

**Aufklärung der Struktur des pho A-Gens aus Escherichia coli
unter besonderer Berücksichtigung seiner Eignung als Vektor**

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
des
Fachbereiches Biologie (23)
der
Freien Universität Berlin

vorgelegt von
Marianne Simonis
aus Karl/Rheinland Pfalz

Berlin 1983



Aufklärung der Struktur des pho A-Gens aus Esche-
richia coli unter besonderer Berücksichtigung
seiner Eignung als Vektor

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
des
Fachbereiches Biologie (23)
der
Freien Universität Berlin

vorgelegt von
Marianne Simonis
aus Karl/Rheinland Pfalz

Berlin 1983

1. Gutachter: Professor Dr. E. R. Lochmann

2. Gutachter: Professor Dr. Dr. E. Gerhards

Tag der mündlichen Prüfung: 9. Juni 1983

Diese Arbeit wurde in den Forschungslaboratorien der Schering AG, Berlin angefertigt.

Herrn Prof. Dr. Dr. E. Gerhards, Herrn Dr. K. Lübke und Herrn Dr. E. Schillinger bin ich für die Förderung und die hervorragenden Arbeitsmöglichkeiten zu Dank verpflichtet. Mein besonderer Dank gilt Herrn Prof. Dr. E.R. Lochmann für die Aufgabenstellung und Herrn Dr. M. Töpert für die ständige Betreuung der Arbeit, für wertvolle Diskussionsbeiträge und Anregungen sowie seine stete Gesprächsbereitschaft. Herrn Dr. G. Siewert danke ich für zahlreiche Diskussionen und Hinweise. Ich danke allen Mitarbeitern der Abteilung Biochemische Pharmakologie, insbesondere Frau G. Norbisch und Frau L. Dummann für die gute Zusammenarbeit.



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41546605_0013

Inhaltsverzeichnis

1.	<u>Einleitung</u>	1
2.	<u>Material und Methoden</u>	8
2.1	Substanzen	8
2.2	Präparation der Plasmide	10
2.2.1	Kulturmedium, Lösungen und Geräte für die Aufarbeitung	10
2.2.2	Versuchsdurchführung	11
2.3	DNA-Isolierung	13
2.3.1	Lösungen und Geräte	14
2.3.2	Spaltung der DNA mit Restriktionsendonucleasen	16
2.3.3	Gelelektrophorese	17
2.3.4	Isolierung der DNA aus Agarosegelen	18
2.3.5	Isolierung der DNA aus Polyacrylamidgelen	20
2.3.6	Alkoholische Fällung	21
2.4	Markierung der DNA mit (^{32}P)	22
2.4.1	Lösungen und Geräte	23
2.4.2	Dephosphorylierung der Restriktionsfragmente	23
2.4.3	Markierung der 5'-DNA-Enden mit Polynucleotidkinase und (γ - ^{32}P)-ATP	24
2.4.4	Markierung der 3'-DNA-Enden mit Deoxynucleotidyltransferase und (α - ^{32}P)Cordycepin 5'-triphosphat	25
2.4.5	Autoradiographie	25
2.5	Kartierung von Restriktionsendonuclease-Schnittstellen	26
2.5.1	Vollständige Spaltung der DNA mit Restriktionsenzymen	27

2.5.2	Partielle Spaltung der DNA mit Restriktionsenzymen	28
2.5.3	Eichkurve und Berechnung der Größe der Restriktionsfragmente	29
2.6	Sequenzierung	29
2.6.1	Lösungen und Geräte	31
2.6.2	Strangtrennung	32
2.6.3	Bestimmung der Nucleotidsequenz	34
2.6.4	Polyacrylamidgel-Elektrophorese	38
2.6.5	Auswertung des Bandenmusters	39
3.	<u>Ergebnisse</u>	41
3.1	Kartierung von Restriktionsendonuclease-Schnittstellen im pho A-Gen	41
3.2	Trennung der Komplementärstränge eines DNA-Fragmentes	48
3.3	Sequenzierungsstrategie	50
3.4	Bestimmung der Nucleotidsequenz des pho A-Gens	53
3.4.1	Schwierigkeiten bei der Interpretation des Bandenmusters	59
3.4.2	Promotor	62
3.4.3	Bindungsstelle für Ribosomen	63
3.4.4	Terminator	64
3.5	Aminosäuresequenz	65
3.6	Analyse des pho A-Gens	68
3.6.1	Erkennungssequenzen für Restriktionsendonucleasen	68
3.6.2	Basenverteilung	71
3.6.3	Häufigkeit des Auftretens von Codewörtern	73
3.6.4	Invertiert repetierte Sequenzen	76

4.	<u>Diskussion</u>	78
4.1	Restriktionskarte	78
4.2	Promotor	81
4.3	Terminator	83
4.4	Aminosäuresequenz	84
4.5	Folgerung	87
5.	<u>Zusammenfassung</u>	89
	<u>Literaturverzeichnis</u>	93

Symbole und Abkürzungen

Abb.	= Abbildung
ATP	= Adenosin-5'-triphosphat
BP	= Basenpaare
DMS	= Dimethylsulfat
DMSO	= Dimethylsulfoxyd
DNA	= Desoxyribonucleinsäure
DTT	= Dithiothreitol
EDTA	= Ethylendiamintetraessigsäure Dinatriumsalz
mRNA	= messenger-RNA
RNA	= Ribonucleinsäure
rpm	= Umdrehungen pro Minute
SDS	= Natriumdodecylsulfat
Tab.	= Tabelle
TEMED	= N, N, N', N'-Tetramethylethylendiamin
Tris	= Tris(hydroxymethyl)-aminomethan
tRNA	= transfer-RNA
U	= Unit
UV-Licht	= ultraviolettes Licht

1. Einleitung

In den sechziger Jahren arbeitete der Schweizer Genetiker Werner Arber an einem damals wenig beachteten Phänomen: der Fähigkeit einiger Bakterienstämme, eingedrungene Fremd-DNA zu zerstören und damit deren Replikation oder gar deren Einbau in das Genom der Wirtszelle zu verhindern (1). Die Restriktion besteht in der Wirkungsweise bis dahin unbekannter Enzyme, die artfremde DNA an spezifischen Nucleotidsequenzen erkennen und sie an dieser Stelle spalten können. 1970 wurde von Smith und Wilcox (2) das erste Restriktionsenzym aus *Haemophilus influenzae* isoliert und charakterisiert. In den folgenden Jahren gelang in rascher Folge die Isolierung einer Vielzahl solcher Restriktionsendonucleasen mit verschiedenen Sequenzspezifitäten, so daß sich nunmehr die Möglichkeit eröffnete, die DNA in definierter Weise zu fragmentieren. Unerwünschtes genetisches Material konnte eliminiert und andersartiges eingefügt werden, denn mit Hilfe der DNA-Ligasen ließen sich die Bruchstücke *in vitro* kovalent miteinander verbinden (3). Damit waren die technischen Voraussetzungen geschaffen für ein Verfahren, das wegen seiner spektakulären Perspektiven nicht nur in Fachjournalen, sondern selbst in Tageszeitungen als "genetische Manipulation", "Genchirurgie" oder "genetic engineering" diskutiert worden ist. Schickte sich die Molekularbiologie doch erstmalig an, die natürliche Barriere zwischen den Arten zu überwinden und so in grundlegende evolutionäre Mechanismen einzugreifen, indem auch Gene von Pro- und Eukaryonten *in vitro* neu kombiniert und in einer Bakterienzelle vermehrt wurden (4, 5).



Solche echten Übertragungen von genetischem Material zwischen sexuell nicht kompatiblen Organismen sind möglich geworden durch die Entdeckung zirkulärer, extrachromosomaler DNA-Moleküle in Prokaryonten und auch in niederen Eukaryonten (6,7). Diese sogenannten Plasmide sind geeignet, heterologe DNA in einer neuen Wirtszelle stabil zu etablieren, da ihre Fähigkeit zur autonomen Replikation auch durch die Erweiterung um die fremde DNA nicht verloren geht. Sie lassen sich leicht isolieren und können mit Hilfe eines Restriktionsenzyms linearisiert werden. Nach Ligieren im Beisein des heterologen chromosomalen DNA-Segments entstehen Hybridplasmide, die in die aufnahmefähig gemachten Bakterienzellen eingeschleust werden. Die Selektion der Zellen, die durch Transformation ein solches Plasmid erhalten haben, wird vereinfacht, wenn dieses die genetische Information für ein charakteristisches biochemisches Merkmal besitzt, für das ein leichter mikrobiologischer Nachweis möglich ist. Am besten geeignet sind die auf den Plasmiden codierten Antibiotikaresistenzen. Diese erlauben nur der transformierten Wirtszelle das Überleben auf Nährböden mit dem entsprechenden Antibiotikum, während die Bakterien ohne Plasmide sensitiv sind und keine Kolonien bilden.

Mit dieser DNA-Klonierungstechnik lassen sich eukaryontische Gene aufgrund der kurzen Generationszeiten von Bakterien rascher vervielfältigen und rein darstellen. Dadurch konnten wichtige Fragen zur Struktur und Funktion des genetischen Materials von höheren Organismen beantwortet werden, was mit den bisherigen klassischen Untersuchungsmethoden nicht möglich war. Doch nicht allein im Erkenntnisgewinn liegt das Ziel der

Gentechnologie, sondern vielmehr noch in der Anwendung zur mikrobiologischen Produktion von Stoffen, die nur in geringer Menge zur Verfügung stehen, wie beispielsweise Insulin und Interferon. Die Gene beider Verbindungen konnten kloniert und auch exprimiert werden (8, 9). Das nach diesem Verfahren hergestellte Insulin ist sogar inzwischen auf den Markt gebracht worden. Erstmals gelang die Expression eines heterologen Gens 1977 mit dem Peptidhormon Somatostatin (10); mittlerweile liegen zahlreiche Berichte über die erfolgreiche Klonierung anderer Gene vor (11,12,13).

Die Mikroorganismen synthetisieren den gewünschten Eiweißstoff jedoch häufig nur in kleinen Mengen, so daß die Produktion wirtschaftlich nicht sinnvoll erscheint. Um die Substanz gentechnisch in ausreichend hoher Konzentration herzustellen, ist es notwendig, den Plasmidvektor dahingehend zu optimieren, daß eine gesteigerte Produktausbeute erwartet werden kann.

Dies gelingt unter anderem durch die Verwendung solcher Plasmide, die in hoher Kopienzahl in der Zelle vorliegen, denn die Menge an Genprodukt ist der Zahl der Genkopien direkt proportional. Insbesondere Plasmide mit niedrigem Molekulargewicht werden diesen Anforderungen gerecht. Sie zeigen zudem eine höhere Transformationsrate. Wegen ihrer Stabilität stellen *E. coli* und seine Plasmide den meist benutzten Typ des Wirts-Vektorsystems dar. Dabei werden die aus dem Plasmid Col EI abgeleiteten Derivate pBR 322 und pBR 325 am häufigsten eingesetzt, da ihre Kopienzahl in Anwesenheit von Chloramphenicol drastisch erhöht wird. In letzter Zeit hat auch das grampositive Bakterium *Bacillus subtilis* als alternativer Wirt für Genklonierungen

Beachtung gefunden (14). Die Hybridplasmide sind in diesem Organismus jedoch weniger stabil als in *E. coli* und werden zuweilen nicht exprimiert (15, 16). Es stehen auch Plasmide der Hefe *Saccharomyces cerevisiae* als Kloniervehikel zur Verfügung (17). Dieser Pilz ist ein einfacher eukaryontischer Organismus, dessen Vorteile darin bestehen, daß er keinerlei humanpathogene Eigenschaften besitzt und sich sehr leicht züchten läßt.

Um die Bakterienzelle zum Produzenten der gewünschten Substanz zu machen, ist es notwendig, die Fremd-DNA hinter dem Promotor des Vektors in das richtige Leseraster des bakteriellen Strukturgens einzufügen, so daß ihre Transkription und Translation dem Bakterien-eigenen Kontrollsystem untersteht. Damit ergibt sich eine weitere Möglichkeit, die Ausbeute des Genproduktes quantitativ zu verbessern, indem die prokaryontischen Regulationselemente solcher Gene verwendet werden, deren Synthese sich induzieren läßt. Dies führt zur verstärkten Transkription der inserierten Gene und damit auch zur Erhöhung der Synthese ihrer Produkte. Bisher wurde meist die Operator/Promotor-Region des β -Galactosidase Gens zur Klonierung eukaryontischer Gene in *E. coli* ausgenutzt. β -Galactosidase ist ein cytoplasmatisches Enzym, dessen Produktion durch die Zugabe von β -Galactosiden angeregt werden kann. Auch die Synthese der alkalischen Phosphatase läßt sich induzieren, wenn die Phosphatkonzentration im Medium niedrig gehalten wird. Dieses Polypeptid ist aber wie die β -Lactamase ein periplasmatisches Enzym und somit leicht zu isolieren, indem die Zellen durch osmotischen Schock aufgebrochen werden. Da die alkalische Phosphatase die Merkmale sowohl der β -Ga-

lactosidase als auch der β -Lactamase trägt, erscheint es vorteilhaft, Plasmidvektoren zu konstruieren, die die der Transkription und Translation dienenden Kontrollabschnitte des *pho A*-Gens in definierten Positionen enthalten, so daß die inserierte Fremd-DNA unter ihrem Einfluß mit großer Effizienz exprimiert werden kann. Zu diesem Zweck wurde von Siewert et al. (18) das die alkalische Phosphatase codierende Segment des *E. coli* Chromosoms in ein Derivat des Plasmids pBR 322 inkorporiert. Ähnliche Plasmide sind von Yoda et al. (19) sowie Inouye et al. (20) konstruiert worden. Das Plasmid pSB 53 (Abb.1) enthält neben dem *pho A*-Gen das *pro C*-Gen, das zu den drei Genen gehört, die für die Biosynthese des L-Prolins aus L-Glutaminsäure notwendig sind.

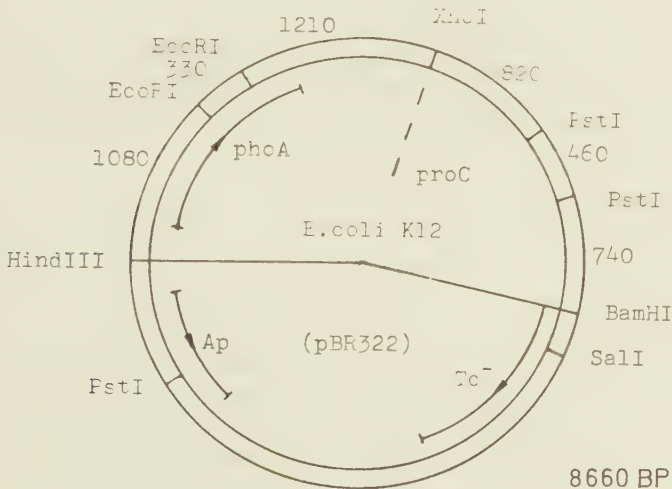


Abbildung 1: Plasmid pSB 53
Die Lage des *pho A*-Gens innerhalb des Abschnittes Hind III-Xho I ist markiert. Das Gen wird in Pfeilrichtung transkribiert.

Ziel der Arbeit ist es, die exakte Position des pho A-Gens innerhalb des Plasmids durch Nucleinsäuresequenzanalyse festzulegen. Damit sollten sich der Promotor und eine DNA-Sequenz identifizieren lassen, deren Transkript als Ribosomenbindungsstelle dienen kann. Erst die Kenntnis über die Lage dieser die Genexpression regulierenden Abschnitte des pho A-Gens ermöglicht die Konstruktion von Plasmidvektoren, die zur Produktion von Hybridproteinen eingesetzt werden können.

In Vorversuchen konnten Siewert et al. (18) bereits feststellen, daß das pho A-Gen auf dem Hind III-Xho I-Fragment des Plasmids pSB 53 liegt und beide Eco RI-Schnittstellen einschließt. Dabei wurde der Promotor in dem Abschnitt Hind III-Eco RI lokalisiert. Das würde bedeuten, daß das Gen innerhalb des Plasmids im Uhrzeigersinn transkribiert wird. Die Leserichtung entspricht damit nicht der in der Genkarte des E. coli Chromosoms angegebenen (21). Sie ließe sich jedoch eindeutig aus der Nucleotidsequenz ableiten.

Weiterhin könnte sich aus dieser Arbeit eine Bestätigung für die Hypothese ergeben, daß auch die alkalische Phosphatase in einer Form synthetisiert wird, die am aminoterminalen Ende um etwa 20 Aminosäuren länger ist als die fertige Polypeptidkette (22). Für alle sekretorischen Proteine wurde ein solches Signalpeptid postuliert, das die Bildung von Kanälen in der Membran induziert, so daß die wachsende Polypeptidkette hindurchtransportiert werden kann (23).

Da diese Aminosäuren während oder nach ihrem Durchtritt mittels einer Signalpeptidase abgetrennt werden, bietet sich - mindestens theore-

tisch - eine Möglichkeit zur Spaltung des aus aminoterminalen bakteriellen Teilstück und gewünschtem eukaryontischen Genprodukt bestehenden Fusionsproteins. Wird das Säugergen hinter einer solchen prokaryontischen Signalsequenz in den Plasmidvektor eingefügt, so wird der Bakterieneigene Proteinanteil möglicherweise enzymatisch entfernt, so daß in dem Periplasma nur das dem Fremdgen entsprechende Produkt vorliegt und isoliert werden kann. Die Insertion von Fremd-DNA hinter der Signalsequenz der alkalischen Phosphatase setzt jedoch voraus, daß sich an der Nahtstelle zwischen Signalpeptid und Strukturprotein codierender Nucleotidsequenz eine Erkennungssequenz für ein Restriktionsenzym befindet. Die Art und Lage der Schnittstelle sind die Grundlage für die Konstruktionspläne der "spacer" und "linker", mit denen die Enden der Fremd- und Vektor-DNA verbunden werden können. Mit Hilfe der Nucleotidsequenz des pho A-Gens sollte sich dieses Problem klären lassen.

2. Material und Methoden

2.1 Substanzen

Alle verwendeten, nachstehend nicht aufgeführten Substanzen wurden in p.a. Qualität von bekannten Firmen bezogen.

Acrylamid von Bio-Rad, Richmond, Kalifornien, USA
alkalische Phosphatase (aus E. coli) von Worthington, Freehold, New Jersey, USA

Agarose (Sea Plaque) von Paesel, Frankfurt
Ammoniumperoxodisulfat von Serva, Heidelberg
(γ -³²P)-ATP von NEN, Boston, Massachusetts, USA

Brij 58 von Atlas Chemie, Essen

Bromphenolblau von Serva, Heidelberg

Caseinhydrolysat (säurehydrolysiert) von Oxoid, London, England

Diethylaminoethylcellulose (DE 52) von Whatmann, Springfield Mill, England

Dimethylsulfat von Ega-Chemie, Steinheim

3'End Labelling System von NEN, Boston, Massachusetts, USA

Ethidiumbromid von Serva, Heidelberg

Formamid von Fluka, Neu-Ulm

Harnstoff (ultrapur) von BRL, Rockville, Maryland, USA

Hydrazin (95%) von Serva, Heidelberg

Hydroxylapathit von Bio-Rad, Richmond, Kalifornien, USA

Lysozym von Serva, Heidelberg

Methylenbisacrylamid von Serva, Heidelberg

Natriumdeoxycholat von Merck, Darmstadt

Natriumkakodylat von BDH, Poole, England

Piperidin von Merck, Darmstadt

Polyethylenglykol (PEG 6000), von Serva, Heidelberg

Proteinase K (aus Pilzen) von Boehringer, Mannheim
Restriktionsendonucleasen zum größten Teil von
BRL, Rockville, Maryland, USA;
sonst von Biolabs, Beverly, Massachusetts, USA
RNase (aus Rinderpankreas) von Boehringer, Mannheim
Röntgenfilme (X-Omat AR 5) von Kodak, Rochester,
USA
Tetramethylethyldiamin (TEMED) von Serva, Heidelberg
T4-Polynucleotidkinase von P+L Biochemicals,
Milwaukee, Wisconsin, USA
tRNA (aus Hefe) von Schwarz/Mann, Heidelberg
Ultrogel ACA 202 von LKB, Bromma, Schweden

2.2 Präparation der Plasmide

Die Plasmid-DNA wurde aus *Escherichia coli* in Anlehnung an das Verfahren von Clevell und Helinski (24) sowie Humphreys et al. (25) isoliert. Die Methode basiert auf einer Verdauung des Mureinsacculus der Bakterienzelle mit Lysozym in Verbindung mit Ethylendiamintetraessigsäure (EDTA) zur Komplexbildung der an die Zelloberfläche gebundenen Metallionen. Es entstehen Sphäroplasten, die durch Zugabe nicht-ionischer Detergenzien wie Brij 58 und Natriumdeoxycholat zur vollständigen Lysis gebracht werden. Durch anschließende Zentrifugation wird die chromosomale DNA präzipitiert. Der Plasmid-haltige Überstand, das Klarlysat, wird durch Cäsiumchlorid-Gradientenzentrifugation gereinigt.

2.2.1 Kulturmedium, Lösungen und Geräte für die Aufarbeitung

<u>Medium:</u>	0,7	%	Na ₂ HPO ₄
	0,3	%	KH ₂ PO ₄
	0,1	%	NH ₄ Cl
	0,25	%	Caseinhydrolysat (säurehydrolysiert)
	2,0	%	Glucose
	1,0	mMol/l	MgSO ₄
	1,0	mg/ml	Vitamin B ₁

<u>Lösung A:</u>	25,0	% (w/v)	Saccharose in
	0,05	Mol/l	Tris-HCl, pH 8
<u>Lösung B:</u>	5	mg/ml	Lysozym in
	0,25	Mol/l	Tris-HCl, pH 8
<u>Lösung C:</u>	0,25	Mol/l	EDTA, pH 8
<u>Lösung D:</u>	1,0	g	Brij 58
	0,4	g	Na-deoxycholat
	25,0	ml	0,25 Mol/l EDTA, pH 8
	20,0	ml	0,25 Mol/l Tris-HCl,
			pH 8
	ad 100 ml mit Aqua bidest		
<u>TES-Puffer:</u>	0,3	% (w/v)	NaCl
	0,05	Mol/l	Tris-HCl, pH 8
	0,05	Mol/l	EDTA
<u>TE-Puffer:</u>	45,0	mMol/l	Tris-HCl, pH 7,9
	0,1	mMol/l	EDTA
<u>Zentrifugen:</u>	Sorvall RC 2-B, Rotor SS 34 und GSA		
	Beckmann Ultrazentrifuge LS-65 oder		
	Sorvall Ultrazentrifuge OTD-2, Rotor		
	50 Ti oder 60 Ti		

2.2.2 Versuchsdurchführung

Eine bei 37°C wachsende Kultur von E.coli K 12 (Stamm HB101) wurde während der log-Phase durch zehnmünütige Zentrifugation bei 8000 g geerntet

(Sorvall, GSA-Rotor, 40°C). Für den folgenden Zellaufschluß, wobei sich die Angaben auf jeweils 1 g Feuchtgewicht beziehen, wurden die Zellen in 8 ml Lösung A suspendiert und 1,6 ml einer frisch bereiteten Lösung B zugegeben. Nach fünf Minuten wurden 3,2 ml Lösung C hinzugefügt und weitere fünf Minuten stengelassen, um dann mit 12,8 ml Lösung D vermischt zu werden. Alle Reaktionsschritte erfolgten im Eisbad. Das Ende der Zelllysis konnte durch eine deutliche Zunahme der Viskosität der Lösung festgestellt werden (10 bis 20 Minuten). Durch anschließende Zentrifugation bei 48.000 g für 30 Minuten entstand als Überstand das Klarlysate (Sorvall, SS 34, 40°C).

Zum Abbau von RNA wurde das Klarlysate 60 Minuten bei 37°C mit RNase (50 µg/ml) inkubiert. Das Enzym wurde vor der Zugabe in wenig TES-Puffer gelöst und 10 Minuten in kochendem Wasser erhitzt. Proteine wurden durch einstündige Behandlung mit Proteinase K (50 µg/ml) bei 37°C abgebaut. Nach Abkühlung wurde die DNA aus dem Klarlysate mit 1/9 Volumenteil einer 3,2 molaren Natriumchloridlösung und 1/4 Volumenteil einer 50%igen Polyethylenglykol-Lösung über Nacht bei 40°C ausgefällt. Die DNA wurde durch zehnminütige Zentrifugation bei 3000 g sedimentiert (Sorvall, SS 34, 40°C), in wenig TES-Puffer aufgenommen und durch zweimalige phenolische Extraktion von Proteinen gereinigt. Durch Zugabe von 2 Volumenteilen Ethanol wurde die DNA bei -18°C über Nacht präzipitiert. Nach Zentrifugation bei 12.000 g und 0°C für 20 Minuten wurde das Sediment getrocknet. Die nachfolgende Cäsiumchlorid-Gradientenzentrifugation bei 100.000 g und 120°C für 40 Stunden trennte zirkuläre von linearer DNA (Beckmann oder Sorvall Ultrazentrifuge, 50 Ti- oder 60 Ti-Rotor).

Im UV-Licht konnten zwei Banden lokalisiert werden. Die Bande mit der höheren Dichte, die die Plasmid-DNA enthält, wurde entnommen. Mit einer Aqua bidest- und Cäsiumchlorid-gesättigten Isopropanol-Lösung wurde das Ethidiumbromid aus der Plasmidpräparation extrahiert. Nach einer anschließenden Dialyse der DNA gegen TE-Puffer wurde die DNA durch Ethanol/Natriumacetat ausgefällt.

2.3 DNA-Isolierung

Plasmide wie auch andere DNA-Moleküle lassen sich mit Hilfe von Restriktionsendonucleasen an den für diese Enzyme charakteristischen Erkennungssequenzen spalten. Die Spaltprodukte können durch Elektrophorese in Agarose- oder Polyacrylamidgelen auf einfache Weise effektiv getrennt werden. Unter dem Einfluß des elektrischen Feldes wandern die negativ geladenen DNA-Moleküle zur Anode, und zwar mit einer Geschwindigkeit, die umgekehrt proportional ist zum Logarithmus der Molekülgröße (26). Durch Zusatz des intercalierenden Fluoreszenzfarbstoffes Ethidiumbromid lassen sich die elektrophoretisch aufgetrennten Restriktionsfragmente durch Bestrahlung mit ultravioletterem Licht sichtbar machen. Diese werden ausgeschnitten und die DNA aus der Gelmatrix isoliert.

2.3.1 Lösungen und Geräte

3-fach konzentrierter Spaltpuffer für die Enzyme Hind III und Xho I:

24	mMol/l	Tris-HCl, pH 7,4
20	mMol/l	MgCl ₂
300	mMol/l	NaCl
20	mMol/l	Mercaptoethanol

Elektrophorese-Puffer:

für Agarosegele:

36	mMol/l	Tris-HCl, pH 7,5
40	mMol/l	NaH ₂ PO ₄
10	mMol/l	EDTA

für Polyacrylamidgele:

50	mMol/l	Tris-Borat, pH 8,3
1	mMol/l	EDTA

Bromphenolblau-Lösung:

0,1 %	(w/v)	Bromphenolblau
100	mMol/l	EDTA
50	% (w/v)	Sucrose

Xylencyanol-Lösung:

0,1 %	(w/v)	Xylencyanol
100	mMol/l	EDTA
50	% (w/v)	Sucrose

Stammlösung für Polyacrylamidgele:

29	% (w/v)	Acrylamid
1	% (w/v)	Methylenbisacrylamid

Kaliumjodid-Lösung:

10	mMol/l	Na ₂ S ₂ O ₅
50	mMol/l	Tris-HCl, pH 7,5
7	Mol/l	KJ

Kaliumphosphat-Puffer:

1	Mol/l	KH ₂ PO ₄ , pH 7
---	-------	--

Tris-EDTA-Puffer:

50	mMol/l	Tris-HCl, pH 7,5
0,1	mMol/l	EDTA

Ethidiumbromid-Lösung:

1	mg/ml	Ethidiumbromid in Ethanol
---	-------	------------------------------

Ammoniumacetat-Lösung:

0,5	Mol/l	NH ₄ -Acetat
-----	-------	-------------------------

Isolierpuffer:

0,5	mMol/l	NH ₄ -Acetat
10	mMol/l	Mg-Acetat
1	mMol/l	EDTA
0,1	% (w/v)	SDS

TE-Puffer:

20 mMol/l Tris-HCl, pH 8
2,5 mMol/l EDTA

Beckmann Mikrofuge B

Zentrifuge: Sorvall RC 2-B, Rotor HB 4 und SS 34

2.3.2 Spaltung der DNA mit Restriktionsendonucleasen

Um das Hind III-Xho I-Fragment, auf dem das pho A-Gen liegt, aus dem Plasmid pSB 53 zu gewinnen, wurden restriktionsendonucleolytische Spaltungen im präparativen Maßstab durchgeführt.

Die in TE-Puffer gelöste Plasmid-DNA wurde mit den Nucleasen Hind III und Xho I versetzt und mit einem den Erfordernissen der Enzyme entsprechenden Puffer vermischt. Es wurde die zwei- bis dreifache vom Hersteller empfohlene Enzymmenge eingesetzt, so daß eine vollständige Verdauung bei optimaler Reaktionstemperatur sicher zu erreichen war. Dies konnte durch analytische Gelelektrophorese überprüft werden.

Nach drei bis fünf Stunden wurden die Enzyme durch zehnminütige Inkubation bei 65°C inaktiviert. Die Probe wurde anschließend in Eis gestellt und kurz zentrifugiert (Beckmann Mikrofuge), um das Kondenswasser an der Gefäßwand zu entfernen. Vor Applikation auf das Gel wurde die Dichte der Auftragslösung durch Zugabe von Saccharose erhöht, so daß eine Unterschichtung des Elektrophorese-Puffers möglich war. Um den Verlauf der Elektro-

phorese beobachten zu können, wurde eine Sucrose-haltige Farbstofflösung auf das Gel gegeben.

Zur Präparation von DNA-Fragmenten für die Nucleotidsequenzbestimmung wurde das Fragment Hind III-Xho I mit einer Reihe von geeigneten Restriktionsenzymen gespalten. Die Versuchsdurchführung entsprach den oben genannten Bedingungen.

2.3.3 Gelelektrophorese

Die Auftrennung der bei der Restriktionsspaltung entstandenen Fragmente erfolgte elektrophoretisch über Agarose- oder Polyacrylamidgele. Die Wahl eines geeigneten Gels ist abhängig von der Größe der Spaltprodukte. Für Fragmente mit 10 bis 1000 Basenpaaren besitzen Polyacrylamidgele das größte Auflösungsvermögen, während sich zur Trennung größerer Fragmente Agarosegele am besten eignen (27).

Das Gießen der Gele (2 mm dick, 16,5 cm breit, 21,5 cm lang) und die Elektrophorese selbst wurden in vertikal orientierten Apparaturen vorgenommen. Für Agarosegele wurden sandgestrahlte Glasplatten verwendet, die eine bessere Haftung der Gelmatrix bewirken. Die Agarosekonzentrationen betrugen 0,8 g, 1 g, 1,2 g und 1,4 g je 100 ml Elektrophorese-Puffer. Die Auftragsvolumina erreichten bei analytischen Ansätzen maximal 20 µl und bei präparativen Versuchsanordnungen höchstens 150 µl pro Geltasche. Dabei wurden nicht mehr als 3 µg beziehungsweise 100 µg DNA aufgetragen, um die Qualität der Trennung der Banden nicht zu gefährden.

Die Agarosegel-Elektrophorese wurde bei einer Spannung von 100 V und einem Stromfluß von 80 mA in ungefähr drei Stunden durchgeführt, wobei ein Ventilator hinter dem Gel aufgestellt worden war, um ein Reißen der Matrix durch Temperaturanstieg zu verhindern. Bei Polyacrylamidgelen hingegen war eine Erwärmung erwünscht, denn dadurch wurde ein schärferes Bandenmuster hervorgerufen. Für diese Gele wurden glatte Glasplatten verwendet. Die Elektrophorese dauerte bei einer konstanten Spannung von 280 V annähernd vier Stunden, wenn Polyacrylamidkonzentrationen von 8% und 10% gewählt wurden. Die Herstellung dieser Gele erfolgte durch entsprechende Verdünnung einer Stammlösung auf 125 ml. Als Initiatoren für die Polymerisation wurden Ammoniumperoxodisulfat und Tetramethylethylendiamin zugesetzt.

2.3.4 Isolierung der DNA aus Agarosegelen

Nach Beendigung der Elektrophorese wurde eine Glasplatte entfernt, das Gel mit Cellulosefolie bedeckt, von der zweiten Glasplatte gelöst und auf eine Kieselgeldünnschichtplatte mit Fluoreszenzindikator gelegt. Da die DNA eine starke Absorption im Bereich von 240 bis 280 nm zeigt, konnten die Restriktionsfragmente bei Bestrahlung mit einer UV-Lampe entsprechender Wellenlänge als fluoreszierende Banden identifiziert werden. Die Banden wurden mit einem Skalpell ausgeschnitten, in 15 ml Corexgläser gegeben und mit einer frisch bereiteten Kaliumjodidlösung versetzt. Dieses chaotrope Agens löst die Gelmatrix auf, ohne die DNA zu beschädigen oder zu denaturieren (28). Die zur vollständigen Auflösung der Agarose erforderliche

Menge an Kaliumjodid richtete sich nach der Agarosekonzentration und der Größe des Gelsegmentes. Eine Inkubation des Reaktionsgemisches bei 37°C beschleunigte den Lösungsprozeß. Unter mäßigem Schütteln löste sich die Agarose in fünf bis zehn Minuten. Anschließend wurde das zweifache Volumen an Aqua bidest zugegeben. Zur Trennung der DNA von der gelösten Agarose wurde eine gesättigte Hydroxylapathit-Lösung hinzugefügt und gut vermischt. Um eine ausreichende Adsorption der DNA an dieses Mineral zu ermöglichen, die durch eine Wechselwirkung der negativ geladenen Phosphatgruppen der Nucleinsäuren mit den Ca^{2+} -Ionen im Gitter des Hydroxylapathits bedingt ist, wurde die Lösung längere Zeit (ca. eine Stunde) bei Raumtemperatur stehengelassen. Durch Zentrifugation bei 16.000 g (Sorvall, HB 4, 4°C) für 15 Minuten wurde die gebundene DNA sedimentiert. Der Überstand wurde verworfen und das Pellet in 600 µl Aqua bidest gelöst und in Eppendorfgefäße überführt. Die Lösung wurde drei Minuten in einer Beckmann Mikrofuge zentrifugiert, der Überstand abgenommen und das Pellet noch zweimal wie beschrieben gewaschen, um das Kaliumjodid vollständig zu entfernen. Das Sediment wurde dann in 100 µl Kaliumphosphat-Puffer aufgeschüttelt. Nach erneuter dreiminütiger Zentrifugation wurde der DNA-haltige Überstand abpipettiert. Dieser Vorgang wurde dreimal wiederholt, so daß eine weitgehende Ablösung der DNA vom Hydroxylapathit erreicht wurde. Die Überstände wurden durch Chromatographie an ACA 202-Gel entsalzt. Zur Elution der DNA mit Tris-EDTA-Puffer wurden Fraktionen à vier Minuten gesammelt. Die Extinktion jeder Fraktion wurde photometrisch bestimmt, indem ein Absorptionsspektrum in einem Bereich von 320 bis 200 nm aufgenommen wurde. Die

damit identifizierten DNA-haltigen Fraktionen wurden vereinigt.

2.3.5 Isolierung der DNA aus Polyacrylamidgelen

Das Polyacrylamidgel wurde in eine Ethidiumbromid-haltige Ammoniumacetat-Lösung gelegt, so daß die elektrophoretisch aufgetrennten Restriktionsfragmente im UV-Licht nach ungefähr 20 Minuten sichtbar wurden und ausgeschnitten werden konnten. Die Extraktion der DNA aus der Gelmatrix erfolgte nach der von Maxam und Gilbert beschriebenen Methode der Diffusion (29).

Der DNA enthaltende Gelstreifen wurde mit einer Skalpellklinge in kleine Würfel zerteilt und diese in eine 5 ml Einwegspritze gefüllt. Der Boden dieser Spritze wurde vorher entfernt und statt dessen ein feinmaschiges Drahtnetz angebracht. Die Gelmatrix wurde durch dieses Sieb in eine 5 ml Einwegpipettenspitze gepreßt, deren Spitze fest verschmolzen und mit silikonisierter Glaswolle ausgestopft worden war. Es wurde 1 ml Isolierpuffer zugegeben, die Pipettenspitze mit Parafilm verschlossen, in ein 15 ml Corexglas eingehängt und über Nacht bei 37°C inkubiert. Aufgrund der hohen Ionenstärke des Puffers diffundierte die DNA leichter aus der Gelmatrix, so daß DNA und Puffer dann durch Filtration durch die Glaswolle von den Gelpartikeln getrennt werden konnten. Zu diesem Zweck wurde die Pipettenspitze geöffnet und die Suspension 10 Minuten bei 3000 rpm in einer Hettich Rotanta/S-Zentrifuge zentrifugiert. Es wurden erneut 500 µl Isolierpuffer zugefügt. Nach sechsstündiger Inkubation bei 37°C folgte eine zweite Zentrifugation unter gleichen Bedingungen.

Das Eluat enthielt lösliche lineare Polyacrylamidmoleküle, die durch die Glaswolle nicht zurückgehalten wurden, so daß eine Reinigung der DNA-Lösung durch DEAE-Cellulose Chromatographie erforderlich war (30). 1 ml Einwegspritzen wurden mit wenig Glaswolle ausgefüllt und mit DE 52-Cellulose beschickt, so daß ein Bettvolumen von 150 μ l erreicht wurde. Die Säulen wurden mit TE-Puffer äquilibriert. Nach Auftragen der DNA-haltigen Lösung wurde die Säule mit 2 ml TE-Puffer gewaschen und die DNA mit 800 μ l 2 molarem Natriumchlorid eluiert. Das Eluat wurde durch Gelchromatographie auf einer 10 ml-Säule (ACA 202) vom Natriumchlorid befreit.

2.3.6 Alkoholische Fällung

Zur Konzentrierung der DNA wurde die wässrige Lösung mit 1/10 Volumenteil 10 molarem Lithiumchlorid und 2 Volumenteilen Ethanol bei -20°C über Nacht oder bei -70°C im Aceton-Trockeneisbad innerhalb von 30 Minuten gefällt. Durch dreißigminütige Zentrifugation bei 23.000 g (Sorvall, HB 4) wurde die DNA bei 4°C sedimentiert. Anschließend wurde die DNA in 1 ml 80%igem Ethanol aufgenommen und das Pellet durch vorsichtiges Drehen des Zentrifugenröhrchens gewaschen, um restliche Salze, die insbesondere die Strangtrennungs- und Sequenzierungsreaktionen stören, zu entfernen. Nach kurzer Fällung im Aceton-Trockeneisbad schloß sich eine zehnminütige Zentrifugation unter den beschriebenen Bedingungen an. Dieser Spülvorgang wurde nochmals wiederholt und das Pellet lyophilisiert.

2.4 Markierung der DNA mit (^{32}P)

Die zur Sequenzbestimmung angewandte Methode von Maxam und Gilbert (29) erfordert eine endständige Markierung der DNA mit dem Isotop ^{32}P . Entsprechend dem für diese Reaktion gewählten Enzym und radioaktiven Substrat können entweder die 5'- oder 3'-Enden eines doppelsträngigen Restriktionsfragmentes phosphoryliert werden.

Bakteriophage T4-induzierte Polynucleotidkinase katalysiert die Übertragung der γ -Phosphatgruppe des ATP an die 5'-Hydroxylenden von Nucleinsäuren (31). Dies setzt jedoch voraus, daß die am 5'-Ende der Polynucleotidkette gebundene Phosphatgruppe vorher enzymatisch abgespalten worden ist.

Zur Reaktion am 3'-Ende wird das Enzym terminale Transferase und in α -Stellung markiertes ^{32}P -Cordycepin-5'-triphosphat (Desoxyadenosin-5'-triphosphat) verwendet. Diese elegante Methode ist von Tu und Cohen (32) beschrieben worden und basiert auf der Tatsache, daß Cordycepin-5'-triphosphat die Synthese eines Oligonucleotides abstoppt, da dieses Nucleotidanalogen keine 3'-Hydroxylgruppe besitzt. Somit kann nur ein einziges Cordycepin-Monophosphat-Molekül angefügt werden. Bei den bisher als Substrat verwendeten (α - ^{32}P)-Ribonucleotiden wurden mehrere Reste am 3'-Ende der DNA angefügt, so daß eine alkalische Hydrolyse und Behandlung mit Phosphatase notwendig war, um einheitlich markierte Fragmente zu erhalten.

2.4.1 Lösungen und Geräte

10-fach konzentrierter Phosphatase-Puffer:

50	mMol/l	Tris-HCl, pH8
5	mMol/l	MgCl ₂

2-fach konzentrierter Kinase-Puffer:

100	mMol/l	Tris-HCl, pH8
10	mMol/l	MgCl ₂
2	mMol/l	DTT

Beckmann Mikrofuge B

2.4.2 Dephosphorylierung der Restriktionsfragmente

Das lyophilisierte Fragment wurde in 300 µl Aqua bidest gelöst, mit 33 µl zehnfach konzentriertem Phosphatase-Puffer sowie 1 µl alkalischer Phosphatase = 1 U versetzt. Das Reaktionsgemisch wurde 40 Minuten bei 60°C inkubiert und anschließend kurz ins Eisbad gestellt. Um die Phosphomonoesterase-Aktivität zu eliminieren, wurde eine phenolische Extraktion durchgeführt. Die Probe wurde mit dem gleichen Volumen Phenol vermischt und zwei Minuten in einer Beckmann Mikrofuge zur Phasentrennung zentrifugiert. Die phenolische Phase wurde verworfen und der Vorgang wiederholt. Die wässrige Phase wurde nacheinander dreimal mit dem doppelten Volumen Diethylether extrahiert, um restliches Phenol zu entfernen.

2.4.3 Markierung der 5'-DNA-Enden mit Polynucleotidkinase und (γ - ^{32}P)-ATP

Die Markierung der 5'-Hydroxylenden der DNA erfolgte stets nach derselben Vorschrift sowohl für glatte als auch kohäsive Strangenden. Da das radioaktive Substrat als wässrige Lösung erhältlich ist, konnte auf das Eindampfen unter Stickstoff verzichtet werden.

Für die Phosphorylierung von 1 pMol DNA wurden standardmäßig 5 pMol = ca. 1,85 MBq (γ - ^{32}P)-ATP eingesetzt. Entsprechend der vorliegenden Mengen an 5'-Hydroxylenden wurde das salzfreie, lyophilisierte Restriktionsfragment in dem benötigten Volumen (γ - ^{32}P)-ATP-Lösung aufgenommen und gelöst. Es wurde das gleiche Volumen an zweifach konzentriertem Kinase-Puffer zugegeben und mit 2 μl = 13 U T4-induzierter Polynucleotidkinase vermischt. Nach dreißigminütiger Inkubation bei 37°C wurden dem Ansatz 12 μg tRNA beigefügt. Das markierte Fragment wurde von dem nicht gebundenen Substrat durch Gelchromatographie auf einer Säule getrennt, die mit ACA 202 gefüllt war. Die markierte DNA wurde aufgrund ihrer Größe als erste eluiert. Die Trennung wurde mittels eines Geiger-Müller-Zählrohres verfolgt, so daß die entsprechende Fraktion aufgefangen werden konnte.

Zur Überprüfung der Markierung wurde ein Aliquot dieser Fraktion auf ein 1,4%iges Agarosegel gegeben und einer Elektrophorese unterworfen. Die Ausbeute wurde durch Messung der Radioaktivität nach Cerenkow bestimmt.

2.4.4 Markierung der 3'-DNA-Enden mit Deoxynucleotidyltransferase und (α - ^{32}P) Cordycepin-5'-triphosphat

Die für diesen Versuch notwendigen Reagenzien wurden einem Kit der Firma NEN entnommen. Auch die Versuchsdurchführung erfolgte in Anlehnung an die vom Hersteller beschriebenen Reaktionsbedingungen (33).

Das salzfreie, lyophilisierte Restriktionsfragment wurde in 8 μl Aqua bidest gelöst und mit 12,5 μl Reaktionspuffer vermischt. Nach Zugabe eines entsprechenden Volumens (α - ^{32}P) Cordycepin-5'-triphosphat wurde der Reaktionsansatz mit 2,5 μl Cobalt(II)chlorid und 3 μl = 18 U Enzym versetzt. Um eine ausreichende Markierung zu erzielen, wurde ein molares Verhältnis (α - ^{32}P) Cordycepin-5'-triphosphat zu 3'-Hydroxylenden von 2:1 gewählt, d.h. daß stöchiometrisch doppelt soviel Substrat wie DNA vorlag. Die Inkubation bei 37°C wurde nach einer Stunde durch Zugabe von 12 μg tRNA beendet. Der Reaktionsansatz wurde zur Trennung der markierten DNA von dem freien (α - ^{32}P) Cordycepin-5'-triphosphat gelchromatographiert (ACA 202). Der weitere Versuchsablauf entsprach dem der 5'-Markierung.

2.4.5 Autoradiographie

Radioaktiv markierte DNA läßt sich in einem Gel autoradiographisch lokalisieren. Zu diesem Zweck wurde die obere Glasplatte nach Abschluß der Elektrophorese von dem Gel entfernt, dieses mit Frappanfolie bedeckt und ein Röntgenfilm aufgelegt. Die Dauer der Exposition richtete sich nach

der Menge der Radioaktivität. Für analytische Zwecke wurde nur wenig Radioaktivität eingesetzt, so daß Expositionszeiten bis zu einer Woche notwendig wurden. Diese konnten durch die Verwendung von Röntgen-Verstärkerfolie verkürzt werden. Bei präparativen Gelen reichte oft eine Minute aus, um die Bande sichtbar zu machen. Bei längeren Expositionszeiten wurden die Gele bei -20°C gelagert, um eine Diffusion der DNA und daraus resultierende unscharfe Bandenmuster zu vermeiden.

3.5 Kartierung von Restriktionsendonuclease-Schnittstellen

Ausgangsmaterial für die angewandte Methode zur Bestimmung der Basensequenz des pho A-Gens sind definierte Restriktionsfragmente (29). Dies erfordert umfangreiche Kenntnisse der auf dem betreffenden DNA-Abschnitt vorkommenden Restriktionsendonuclease-Schnittstellen.

Das einfachste Kartierungsverfahren ist die vollständige Spaltung des Moleküls mit dem betreffenden Enzym. Wird eine Marker-DNA mit bekannten Fragmentgrößen parallel dazu gelelektrophoretisch aufgetrennt, so lassen sich Anzahl und Größe der entstandenen Spaltprodukte bestimmen. Die Anordnung der Schnittstellen kann damit jedoch nicht ermittelt werden. Diese läßt sich nach der von Smith und Birnstiel (34) beschriebenen Methode der partiellen Spaltung einseitig ^{32}P -endmarkierter DNA festlegen. Es entstehen zahlreiche Fragmente, von denen im Autoradiogramm nur die sichtbar sind, die mit dem markierten Ende verbunden bleiben. Die

Länge dieser Bruchstücke entspricht der Entfernung der Schnittstelle vom markierten Ende.

2.5.1 Vollständige Spaltung der DNA mit Restriktionsenzymen

Zur Untersuchung der Restriktionsendonuclease-Schnittstellen in den Fragmenten Hind III-Eco RI und Eco RI-Xho I wurden folgende Enzyme eingesetzt: Alu I, Ava I, Bgl I, Hae II, Hae III, Hha I, Hinc II, Hinf I, Hpa I, Hpa II, Kpn I, Mst I, Nru I, Pvu II, Rsa I, Sau 3A, Sau 96I, Sma I, Sph I, Sst II, Tac I, Taq I, Xba I, Xor II.

Je 1 µg DNA wurde mit der gewünschten Endonuclease und dem entsprechenden Puffer vermischt. Die Inkubation erfolgte nach den Angaben des Herstellers, wobei die benötigte Enzymmenge verdoppelt und die Inkubationszeit erhöht wurde. Nach drei Stunden wurde die Reaktion bei 65°C abgestoppt. Für die Enzyme Tac I und Taq I wurde eine Inaktivierungstemperatur von 75°C gewählt. Jede Probe wurde mit 5 µl Bromphenolblau versetzt und auf ein 10%iges analytisches Polyacrylamidgel aufgetragen. Die Elektrophorese wurde bei 280 V solange durchgeführt, bis der Farbstoff das Gelende erreicht hatte. Durch Anfärben des Gels mit Ethidiumbromid in einer Ammoniumacetat-Lösung wurde das Bandenmuster in kurzweiligem UV-Licht sichtbar und konnte durch Vorschalten eines gelben Filters photographiert werden.

2.5.2 Partielle Spaltung der DNA mit Restriktionsenzymen

Im Gegensatz zur vollständigen Verdauung erfordert diese Methode der partiellen Spaltung einseitig endmarkierte DNA-Fragmente. Die Fragmente Pst I-Eco RI, Hind III-Xho I und Eco RI-Pst I wurden mit ^{32}P 5' markiert. Durch anschließende restriktions-endonucleolytische Spaltung mit Hind III (Pst I-Eco RI), Eco RI (Hind III-Xho I) oder Xho I (Eco RI-Pst I) entstanden Fragmente, die nur an einem 5'-Ende eine radioaktive Phosphatgruppe trugen. Nach elektrophoretischer Auftrennung wurden die Banden Hind III-Eco RI und Eco RI-Xho I isoliert, die entweder am Hind III- oder am Eco RI-Ende markiert waren.

In dem Abschnitt Eco RI-Xho I wurde die Lage der Schnittstellen für die Enzyme Hinf I, Hae II und Sph I bestimmt, während das Fragment Hind III-Eco RI nur partiell gespalten wurde mit den Restriktionsendonucleasen Alu I, Hinf I, Hha I, Hpa II, Hae III, Hinc II, Pvu II, Rsa I und Taq I.

In Vorversuchen wurde zunächst die Enzymkonzentration ermittelt, die das betreffende DNA-Molekül nur unvollständig verdaut. Eine entsprechende Menge wurde der DNA zugesetzt, mit Puffer vermischt und bei einer geeigneten Reaktionstemperatur inkubiert. Nach 0, 5, 15 und 60 Minuten wurden Proben gleichen Volumens entnommen und 10 Minuten bei 65°C beziehungsweise 75°C gehalten, um die Enzyme zu inaktivieren. Die Proben wurden mit je 5 µl Bromphenolblau versetzt und in einem analytischen 10%igen Polyacrylamidgel aufgetrennt. Das entstandene Bandenmuster wurde autoradiographisch dokumentiert.

2.5.3 Eichkurve und Berechnung der Größe der Restriktionsfragmente

Zur Bestimmung der Molekulargewichte der entstandenen Restriktionsfragmente wurde eine DNA-Standardmischung mit bekannten Molekülgrößen geelektrophoretisch aufgetrennt und die Laufstrecken bestimmt, indem der Abstand vom Boden der Geltasche bis zur jeweiligen DNA-Bande gemessen wurde. Die halblogarithmische Darstellung der Molekulargewichte gegenüber der Wanderungsstrecke im Gel ergab eine Gerade. Aus dieser Eichkurve ließ sich die Größe der Restriktionsfragmente durch Interpolation bestimmen. Als Standardmischung wurden die bekannten elf Hae III-Fragmente der DNA des Bakteriophagen ϕ X 174 benutzt.

2.6 Sequenzierung

Die Bestimmung der Nucleotidsequenz des pho A-Gens erfolgte im wesentlichen nach der von Maxam und Gilbert ausgearbeiteten Methode (29). Sie basiert auf einer basenspezifischen Modifikation, gefolgt von einer durch Piperidin induzierten Abspaltung der veränderten Basen von der Pentose und Strangbruch der DNA bei diesem Zuckerrest. Die selektiven hydrolytischen Spaltungen wurden durch die beiden Reagenzien Dimethylsulfat und Hydrazin ausgelöst.

Dimethylsulfat methyliert den Stickstoff des Guanin in Position 7 und des Adenin in Position 3. Diese Modifikation führt zu einer Destabilisierung

der N-glykosidischen Bindung, die beim Erhitzen mit Piperidin zerfällt, wobei bevorzugt die Guaninreste abgespalten werden. Das von Maxam und Gilbert angegebene Verfahren zur Abspaltung der Adeninreste wurde nicht angewendet. Statt dessen wurde die DNA, wie von Burton (35) für die Purinbasen vorgeschlagen, unter schwach sauren Bedingungen hydrolysiert. Dabei werden nur die Purine entfernt, ohne daß die Bindungen zwischen Pyrimidinen und Desoxyribose oder die Phosphodiesterbindungen beeinträchtigt werden. Die resultierende Apurinsäure wird durch nachfolgende Behandlung mit Piperidin gespalten.

Hydrazin reagiert mit den Pyrimidinen Cytosin und Thymin, wobei die Base abgespalten und ein Ribosylharnstoffrest hinterlassen wird. Dieser reagiert mit Hydrazin weiter zu Hydrazon. Durch Zugabe von Piperidin wird der mit dem Zucker veresterte Phosphatrest eliminiert. Die Reaktionsgeschwindigkeit ist für beide Basen gleich, in Anwesenheit von 2 molarem Natriumchlorid wird die Reaktion von Thymin unterdrückt, so daß nur Cytosin chemisch verändert wird und damit die beiden Basen unterschieden werden können.

Die Reaktionsbedingungen sind so gewählt, daß im Mittel nur ein Bruch pro Molekül erfolgt. Das erhaltene Fragmentgemisch wird gelektrophoretisch unter denaturierenden Bedingungen aufgetrennt und die Sequenz vom Autoradiogramm abgelesen.

Da dieses Verfahren von einseitig endmarkierten DNA-Fragmenten ausgeht, müssen die Stränge durch Denaturieren und anschließende Gelektrophorese getrennt werden, so daß sie einzeln zu isolieren sind. Ist diese Methode der Strangtrennung nicht

erfolgreich, so wird mit Hilfe der Restriktionskarte ein zweites Restriktionsenzym ausgewählt, das das Molekül in zwei ungleich lange Teile spaltet. Die Spaltprodukte werden in präparativen Agarose- oder Polyacrylamidgelen elektrophoretisch separiert, autoradiographisch lokalisiert und, wie in Kapitel 2.3.4 und 2.3.5 beschrieben, isoliert.

2.6.1 Lösungen und Geräte

2-fach konzentrierter Strangtrennungspuffer:

30	% (v/v)	DMSO
1	mMol/l	EDTA
0,05	% (w/v)	Xylencyanol
0,05	% (w/v)	Bromphenolblau

Natriumchlorid-Natriumhydroxid-Lösung:

2	Mol/l	NaCl
0,1	Mol/l	NaOH

Natrium-Kakodylat-Puffer:

50	mMol/l	Na-Kakodylat, pH 8
1	mMol/l	EDTA

DMS-Stop-Lösung:

1	Mol/l	Na-Acetat, pH 7
1	Mol/l	Mercaptoethanol
100	µg/ml	tRNA

Apu-Mix:

66	% (v/v)	Ameisensäure
2	% (w/v)	Diphenylamin
1	mMol/l	EDTA

Natriumchlorid-Lösung:

5	Mol/l	NaCl
---	-------	------

Na-Ac-Mix:

0,3	Mol/l	Na-Acetat, pH 7,5
0,1	mMol/l	EDTA

NaOH-Lösung:

5	Mol/l	NaOH
---	-------	------

FeCl₃-Lösung:

3	Mol/l	FeCl ₃
---	-------	-------------------

Zentrifuge: Sorvall RC 2-B, Rotor SS 342.6.2 Strangtrennung

Die Versuchsbedingungen wurden nach den Angaben von Maxam und Gilbert gewählt (29). 1 µg DNA wurde in 20 µl Aqua bidest gelöst und mit 20 µl zweifach konzentriertem Strangtrennungspuffer vermischt. Der Reaktionsansatz wurde zwei Minuten bei 90°C inkubiert, anschließend im Eisbad abgekühlt und

sofort auf ein präparatives 5%iges Polyacrylamid-gel (Acrylamid:Bisacrylamid = 60:1) aufgetragen. Die Elektrophorese wurde über Nacht bei 60 V durchgeführt, um eine Erwärmung und damit eine Renaturierung der DNA-Einzelstränge zu verhindern.

Nach etwa 15 bis 18 Stunden wurde die Elektrophorese beendet. Der Röntgenfilm wurde nur kurz bei Raumtemperatur aufgelegt, um zu verhindern, daß durch Überstrahlen die Trennung der Banden nicht mehr erkennbar war. Die DNA-haltigen Gelstreifen wurden ausgeschnitten. Nach Zerkleinern der Gelstreifen mit Hilfe eines Drahtsiebes wurde die Gelmatrix im Isolierpuffer mazeriert, damit die DNA nach dem beschriebenen Verfahren (Kapitel 2.3.5) durch Filtration von den Gelpartikeln getrennt werden konnte.

Die Reinigung des Filtrats wurde in DE 52-Cellulose Säulen vorgenommen, wobei die DNA mit 800 µl einer Natriumchlorid-Natriumhydroxid-Lösung eluiert wurde (30). Durch Gelchromatographie wurde dieses Eluat entsalzt.

Um festzustellen, welche Restriktionsfragmente erfolgreich einer präparativen Strangtrennung unterworfen werden können, wurden Vorversuche durchgeführt. Nicht markierte DNA wurde nach den beschriebenen Reaktionsbedingungen denaturiert und elektrophoretisiert. Mit Hilfe von Ethidiumbromid wurden die DNA-Banden im Gel unter UV-Licht sichtbar gemacht, so daß erfolgreiche Trennungen von Komplementärsträngen identifiziert werden konnten.

2.6.3 Bestimmung der Nucleotidsequenz

Wie von Maxam und Gilbert gefordert, wurden für die einzelnen Reaktionsschritte sehr reine Substanzen eingesetzt. Dies gilt insbesondere für Dimethylsulfat, Hydrazin und Piperidin. Aus diesem Grunde wurden Aliquots dieser Reagenzien in Eppendorfgefäße gefüllt, fest verschlossen und bei -20°C aufbewahrt. Es wurden nur silikonisierte Reaktionsgefäße benutzt, so daß einem Verlust an DNA durch Adsorption vorgebeugt werden konnte.

Das lyophilisierte, einseitig markierte DNA-Fragment wurde in Aqua bidest gelöst und so auf vier Eppendorfgefäße verteilt, daß jeweils zwei Gefäße doppelt soviel DNA enthielten wie die beiden anderen. Jede Probe wurde mit 6 μg ultrabeschallter Kalbsthymus-DNA versetzt und lyophilisiert. Diese Lösung sowie die Hefe-tRNA waren vor ihrer Verwendung phenolisch extrahiert worden, um Proteine zu entfernen. Alle Lösungen wurden während des Versuchs in Eis gehalten.

Guanin spezifische Reaktion (G)

DNA lösen in	200	μl Natrium-Kakody-
		lat-Puffer
und	0,5	μl DMS zufügen

Inkubation bei 20°C für 2,5 Minuten

im Eisbad zugeben	50	μl DMS-Stop-Lösung
und	750	μl Ethanol

mischen und mindestens 10 Minuten bei -70°C (Aceton-Trockeneisbad) stehen lassen

sofort auf ein präparatives 5%iges Polyacrylamid-gel (Acrylamid:Bisacrylamid = 60:1) aufgetragen. Die Elektrophorese wurde über Nacht bei 60 V durchgeführt, um eine Erwärmung und damit eine Renaturierung der DNA-Einzelstränge zu verhindern.

Nach etwa 15 bis 18 Stunden wurde die Elektrophorese beendet. Der Röntgenfilm wurde nur kurz bei Raumtemperatur aufgelegt, um zu verhindern, daß durch Überstrahlen die Trennung der Banden nicht mehr erkennbar war. Die DNA-haltigen Gelstreifen wurden ausgeschnitten. Nach Zerkleinern der Gelstreifen mit Hilfe eines Drahtsiebes wurde die Gelmatrix im Isolierpuffer mazeriert, damit die DNA nach dem beschriebenen Verfahren (Kapitel 2.3.5) durch Filtration von den Gelpartikeln getrennt werden konnte.

Die Reinigung des Filtrats wurde in DE 52-Cellulose Säulen vorgenommen, wobei die DNA mit 800 µl einer Natriumchlorid-Natriumhydroxid-Lösung eluiert wurde (30). Durch Gelchromatographie wurde dieses Eluat entsalzt.

Um festzustellen, welche Restriktionsfragmente erfolgreich einer präparativen Strangtrennung unterworfen werden können, wurden Vorversuche durchgeführt. Nicht markierte DNA wurde nach den beschriebenen Reaktionsbedingungen denaturiert und elektrophoretisiert. Mit Hilfe von Ethidiumbromid wurden die DNA-Banden im Gel unter UV-Licht sichtbar gemacht, so daß erfolgreiche Trennungen von Komplementärsträngen identifiziert werden konnten.

im Eisbad zugeben	200	µl Na-Ac-Mix
und	750	µl Ethanol

mischen und mindestens 10 Minuten bei -70°C (Aceton-Trockeneisbad) stehen lassen

Die im Aceton-Trockeneisbad gefällten DNA-Proben wurden 15 Minuten bei 4°C und 17.000 g (Sorvall, SS 34) zentrifugiert, die Überstände abgenommen und in die vorbereiteten NaOH- und FeCl_3 -Lösungen gegeben, um das restliche DMS und Hydrazin zu neutralisieren. Die Sedimente wurden in 250 µl 0,3 molarem Na-Acetat aufgenommen und mit 750 µl Ethanol vermischt. Die DNA wurde 10 bis 15 Minuten bei -70°C gefällt. Die anschließende Zentrifugation dauerte 10 Minuten unter den angegebenen Bedingungen. Die Überstände wurden verworfen und der Vorgang wiederholt. Die Präzipitate wurden dann mit 1 ml 80%igem Ethanol gewaschen, fünf Minuten bei -70°C gehalten, fünf Minuten wie beschrieben zentrifugiert und die Überstände wieder abgenommen. Nachdem dieser Waschvorgang nochmals durchgeführt worden war, wurden die Sedimente lyophilisiert.

Piperidin-Spaltung

Die lyophilisierte DNA wurde in 100 µl 1 molarem Piperidin aufgenommen und gelöst. Die Verdünnung des Piperidins erfolgte bei jedem Versuch neu und wurde erst kurz vor Durchführung dieses Reaktionsschrittes vorgenommen. Nach kurzer Zentrifugation der Proben in einer Beckmann Mikrofuge folgte eine

Inkubation bei 90°C (Wärmeschrank) eine Stunde lang. Die Eppendorfggefäße wurden mit einer Metallplatte bedeckt, so daß ein fester Verschuß der Proben bei dieser hohen Temperatur sichergestellt werden konnte. Anschließend wurde das Piperidin im Vakuum abgezogen. Die Proben wurden noch zweimal in Aqua bidest aufgenommen - zunächst in 50 µl, dann in 30 µl - und lyophilisiert.

Eine weitere Methode zur Entfernung der Piperidinlösung, die zur Anwendung kam, besteht in der Fällung der DNA (30). Zu diesem Zweck wurden die Proben in neue Eppendorfggefäße überführt, mit 100 µl 0,3 molarem Na-Acetat und 800 µl Ethanol vermischt und 30 Minuten bei -70°C gehalten. Es schloß sich eine dreißigminütige Zentrifugation bei 4°C und 17.000 g (Sorvall, SS 34) an. Die Überstände wurden abpipettiert und die Sedimente mit 1 ml 80%igem Ethanol gewaschen, fünf Minuten bei -70°C gefällt und fünf Minuten bei der angegebenen Umdrehungszahl zentrifugiert. Nach Lyophilisation wurden die Proben nochmals in 30 µl Aqua bidest aufgenommen und erneut lyophilisiert.

Die erhaltenen Proben wurden in 5 bis 8 µl einer Farbstoff-haltigen Formamid-Lösung aufgenommen und in Abhängigkeit von der Anzahl der Gele auf zwei oder drei Eppendorfggefäße verteilt. Die Proben wurden dann für zwei Minuten in einem Wasserbad bei 90°C inkubiert, im Eisbad abgekühlt und anschließend auf Polyacrylamidgele appliziert.

2.6.4 Polyacrylamidgel-Elektrophorese

Die elektrophoretische Auftrennung der denaturierten DNA-Fragmente erfolgte in Polyacrylamidgelen unterschiedlicher Konzentrationen (36). Mit Hilfe eines 20%igen Gels lassen sich Oligonucleotide mit einer Länge von 2 bis 25 Basen so voneinander trennen, daß ein lesbares Autoradiogramm entsteht. Für längere Bruchstücke eignen sich 8%ige Gele am besten.

20%iges Gel

19,3	% (w/v)	Acrylamid
0,7	% (w/v)	Bisacrylamid
8,3	Mol/l	Harnstoff
0,07	% (w/v)	Ammoniumperoxodisulfat
0,008	% (v/v)	TEMED

8%iges Gel

7,7	% (w/v)	Acrylamid
0,3	% (w/v)	Bisacrylamid
8,3	Mol/l	Harnstoff
0,07	% (w/v)	Ammoniumperoxodisulfat
0,02	% (v/v)	TEMED

Zunächst wurden Acrylamid und Bisacrylamid in wenig Tris-Borat Puffer bei 37°C gelöst und sukzessiv die entsprechende Menge Harnstoff zugegeben. Nach vollständiger Auflösung wurde das gewünschte Volumen eingestellt und die Lösung entgast. Erst dann wurden Ammoniumperoxodisulfat und TEMED zugesetzt.

Die Abmessungen der Gele betrugen 42,5 cm x 35 cm. 8%ige Gele erreichten eine Stärke von 0,3 mm, bei 20%igen Gelen dagegen war die Gelschicht 0,5 mm dick. Vor dem Auftragen der Proben wurden die Gele für etwa zwei bis vier Stunden in einem elektrischen Stromkreis angeschlossen. Während dieser Präelektrophorese sollte der Stromfluß 20 mA nicht überschreiten. Erst bei Erreichen einer Spannung von 2000 V wurden die Ansätze aufgetragen.

Die Elektrophorese wurde im Falle des 20%igen Gels abgebrochen, wenn Bromphenolblau 1/3 der Gellänge passiert hatte. Hatte der Farbstoff das Gelende des ersten 8%igen Gels erreicht, so konnte die Elektrophorese beendet werden. Das zweite 8%ige Gel wurde solange der Elektrophorese unterworfen, bis Xylencyanol 3/4 des Gels durchlaufen hatte. Die Gele wurden mit einem strahlenempfindlichen Film bedeckt und in Gegenwart von Röntgen-Verstärkerfolie bei -20°C exponiert.

2.6.5 Auswertung des Bandenmusters

Die aus den vier verschiedenen Reaktionen hervorgegangenen Produkte partieller Hydrolyse wurden nebeneinander gelektrophoretisch aufgetrennt. Es entstanden Nucleotidsequenz-spezifische Bandenmuster, wie sie in dem nachfolgenden Schema dargestellt sind. Über den Gelbahnen sind die Basenspezifitäten der Spaltungen und rechts die Sequenz angegeben. Diese konnte nach der Autoradiographie direkt vom Film abgelesen werden. Die Reihenfolge der Nucleotide wurde rekonstruiert, indem das kleinste Bruchstück ermittelt und das Nucleotid

angefügt wurde, dessen Zerstörung zur Bildung des Spaltstückes geführt hatte.

G	G+A	C+T	C	Sequenz
---	-----	-----	---	---------

-	-	-	-	G
		-		T
	-			C
	-			A
		-		A
		-	-	T
				C

Die nach diesem Verfahren zusammengestellte komplette Nucleotidsequenz des pho A-Gens wurde mit Hilfe des Computerprogramms NUCDAT (37) unter Anleitung von K. Isono in einer Dec 10-Rechneranlage analysiert.

3. Ergebnisse

3.1 Kartierung von Restriktionsendonuclease-Schnittstellen im pho A-Gen

Die von Maxam und Gilbert beschriebene Methode zur Sequenzbestimmung ist limitiert durch das Auflösungsvermögen der Gele. Mit Hilfe der benutzten etwa dreiundvierzig Zentimeter langen Polyacrylamidgele zweier verschiedener Konzentrationen ist es nicht möglich, mehr als zweihundert Basen gut aufzutrennen. Demzufolge erfordert die Ermittlung der Gesamtsequenz des pho A-Gens die Zerlegung der DNA in Stücke passender Länge, die in der Größenordnung einiger Hundert, mindestens jedoch von zweihundert Nucleotiden liegen sollte. Die Herstellung solcher Fragmente setzt voraus, daß die Restriktionsendonuclease-Schnittstellen auf dem betreffenden DNA-Abschnitt bekannt sind und ihre Anordnung in einer Restriktionskarte festgelegt wird.

Untersuchungen von Siewert et al. (18) haben gezeigt, daß das pho A-Gen im Hind III-Xho I-Fragment des Plasmids pSB 53 lokalisiert ist. Beide Eco R I-Schnittstellen liegen innerhalb des Gens. Ausgehend von diesem Kenntnisstand wurden die Fragmente Hind III-Eco RI und Eco RI-Xho I mit verschiedenen Restriktionsendonucleasen sowohl vollständig als auch partiell gespalten. Die entstandenen Spaltprodukte wurden gelektrophoretisch aufgetrennt und entweder im UV-Licht als fluoreszierende Banden oder autoradiographisch sichtbar gemacht. Ihre Größen konnten durch Vergleichen und Interpolieren der Wanderungsdistanzen

von DNA-Fragmenten bekannter Länge ermittelt werden.

In Tabelle 1a und 1b sind Anzahl und Größe der Spaltstücke angegeben, die bei vollständiger Hydrolyse der Fragmente Hind III-Eco RI und Eco RI-Xho I entstehen. Damit ist zwar der Abstand zwischen den Schnittstellen festgelegt, doch ihre Anordnung im Molekül läßt sich daraus nicht ableiten. Diese resultiert aus den Wanderungstrecken der bei partieller Verdauung ^{32}P -endmarkierter DNA gebildeten Bruchstücke, denn ihre Größe entspricht dem Abstand vom markierten Ende der DNA bis zur jeweiligen Schnittstelle.

Die Auswertung der zugehörigen autoradiographischen Bandenmuster gestaltete sich zuweilen bei der Identifizierung schwacher Banden recht schwierig, da diese nicht unbedingt eine Schnittstelle der DNA mit der ausgewählten Restriktionsendonuclease anzeigen, sondern auch auf Kontaminationen des Enzyms durch andere Endonucleasen zurückzuführen sind. Insbesondere bei Enzymen mit häufigen Erkennungssequenzen wie Hha I und Hpa II im Fragment Hind III-Eco RI gelang eine eindeutige Zuordnung nicht. Durch einen Vergleich mit den aus der vollständigen Spaltung gewonnenen Ergebnissen konnten nachträglich einige Unstimmigkeiten beseitigt werden. Auch die Durchführung der partiellen Spaltungen des Fragmentes Hind III-Eco RI sowohl mit Hind III- als auch mit Eco RI markiertem Ende machte eine weitere Aufklärung möglich.

Tabelle 1a: Berechnete Fragmentgröße und Anzahl der Schnittstellen in dem Hind III-Eco RI-Abschnitt nach vollständiger Spaltung mit Restriktionsendonucleasen

Restriktionsenzym	Fragmentgröße in Basenpaaren	Anzahl der Schnittstellen
Alu I	744	3
	175	
	139	
	71	
Hae II	1005	1
	87	
Hae III	875	1
	220	
Hha I	470	10
	121	
	110	
	98	
	73	
	53	
	36	
	32	
	30	
	10	
	7	
Hinc II	664	1
	409	
Hinf I	439	4
	332	
	135	
	92	
	61	
Hpa II	333	8
	191	
	175	
	98	
	92	
	36	
	22	
	15	
	14	

Pvu II	763	2
	200	
	173	
Rsa I	819	2
	167	
	84	
Sau 3A	360	4
	300	
	170	
	150	
	70	
Sau 96I	859	1
	200	
Taq I	700	1
	365	
Ava I	-	-
Bgl I	-	-
Kpn I	-	-
Sph I	-	-
Xba I	-	-

Tabelle 1b: Berechnete Fragmentgröße und Anzahl
der Schnittstellen in dem Abschnitt
Eco RI-Xho I nach vollständiger Spal-
tung mit Restriktionsendonucleasen

Restriktionsenzym	Fragmentgröße in Basenpaaren	Anzahl der Schnittstellen
Alu I	744	3
	565	
	94	
	37	
Hae II	762	2
	391	
	105	
Hae III	835	6
	156	
	126	
	110	
	103	
	58	
Hinf I	19	1
	1076	
Hha I	167	10
	400	
	258	
	179	
	113	
	85	
	78	
	71	
	69	
	56	
	32	
	21	
Hpa II	664	6
	194	
	121	
	100	
	96	
	92	
	76	

Sau 96I	695	
	171	
	152	5
	121	
	98	
	55	
Sph I	981	1
	200	
Tac I	356	
	333	
	296	
	171	6
	56	
	55	
	29	
Hpa I	-	-
Hinc II	-	-
Mst I	-	-
Nru I	-	-
Pvu II	-	-
Rsa I	-	-
Sma I	-	-
Sst II	-	-
Xor II	-	-

Die in Abbildung 2a und 2b gezeigten Restriktionskarten der Fragmente Hind III-Eco RI und Eco RI-Xho I enthalten die nach diesem Verfahren festgelegten Schnittstellen. Für die Enzyme Alu I, Hinf I, Hae II, Hae III, Hinc II, Pvu II, Rsa I, Sph I und Taq I konnte die Lage ihrer Schnittstellen im Rahmen der durch das graphische Auswertungsverfahren bedingten Fehlergrenzen bestimmt werden, während bei den Restriktionsendonucleasen Hha I und Hpa II aus den genannten Gründen auf eine Eintragung nicht eindeutig festzulegender Schnittstellen verzichtet wurde. Ihre Reihenfolge ergab sich mit dem Fortschreiten der Nucleotidsequenzanalyse.

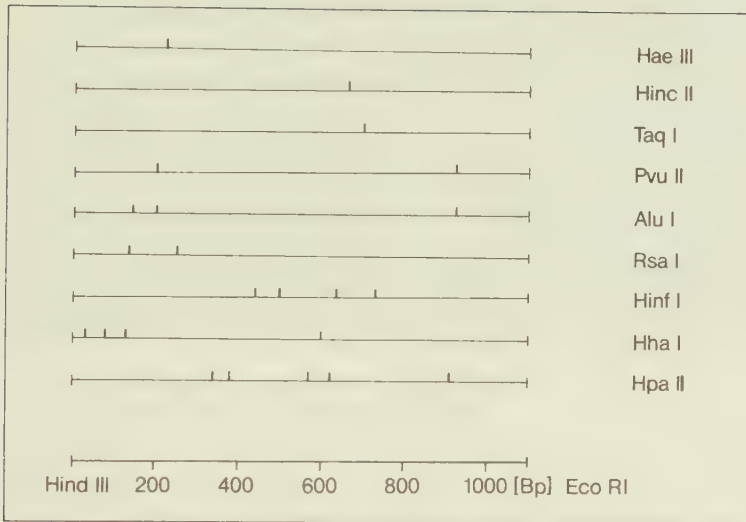


Abbildung 2a: Restriktionskarte des Fragmentes Hind III-Eco RI

Die Lage der verschiedenen Restriktionsendonuclease-Schnittstellen innerhalb des Fragmentes Hind III-Eco RI ist durch Querstriche gekennzeichnet. Die Nucleotide sind von der Hind III-Schnittstelle ausgehend bis zum Eco RI-Ende durchnummeriert, so daß anhand dieser Skala die Länge der Restriktionsfragmente in BP bestimmt werden kann.

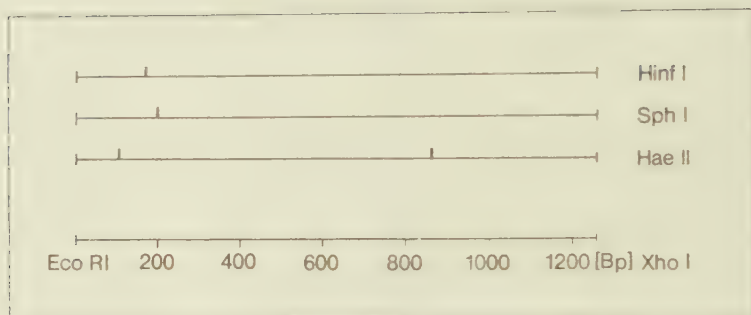


Abbildung 2b: Restriktionskarte des Fragmentes Eco RI-Xho I

Die Querstriche markieren die Lage der Schnittstellen für die Restriktionsenzyme Hae II, Hinf I und Sph I innerhalb des Abschnittes Eco RI-Xho I. Seine Nummerierung beginnt an der Eco RI-Schnittstelle und erfolgt fortlaufend bis zum Xho I-Ende.

3.2 Trennung der Komplementärstränge eines DNA-Fragmentes

Die Zerlegung des doppelsträngigen DNA-Moleküls in seine Einzelstränge durch Denaturierung und anschließende Gelelektrophorese stellt ein geeignetes Verfahren zur Herstellung einseitig endmarkierter DNA dar, wie sie die Sequenzanalyse nach Maxam und Gilbert zwingend vorschreibt (29).



1 2 3 4 5 6

Abbildung 3: Einzelstrangtrennung durch Elektrophorese in einem präparativen Polyacrylamidgel (5%)

Die mit den Ziffern 1 bis 6 gekennzeichneten Gelbahnen repräsentieren folgende Restriktionsfragmente:

- 1 = Sau 3A-Sau 3A (448 - 632),
- 2 = Hinc II-Eco RI (339 - 795),
- 3 = Hinf I-Hinf I (1065 - 1288),
- 4 = Eco RI-Eco RI (795 - 1126),
- 5 = Hinf I-Hinf I (459 - 813).
- 6 = Sau 3A-Sau 3A (148 - 448)

Die in Klammern gesetzten Zahlen kennzeichnen die Lage der Restriktionsendonuclease-Schnittstellen innerhalb des pho A-Gens, wobei die jeweilige Positionsnummer der 1. Base der zugehörigen Erkennungssequenz entspricht, wie in Tabelle 3 angegeben.

Die Gründe, warum die Komplementärstränge gelelektrophoretisch getrennt werden können, sind

nicht vollständig geklärt. Möglicherweise liegt die Ursache in einer durch die Basenzusammensetzung bedingten Faltung jeder Einzelkette nach vollständiger Denaturierung des Doppelstranges (38). Somit könnten die Unterschiede in ihrer Wanderungsgeschwindigkeit auf ungleiche Molekülgestalten zurückgeführt werden. Doch oft sind die Differenzen in den elektrophoretischen Mobilitäten der komplementären Stränge nicht ausreichend, um diese einzeln aus der Gelmatrix isolieren zu können. Die in Abbildung 3 dargestellten Restriktionsfragmente 1 bis 5 lassen sich durch Gelelektrophorese gut sichtbar in ihre Einzelketten zerlegen, während für das Fragment 6 mit dieser Methode keine erkennbare Trennung erreicht wurde.

2.3 Sequenzierungsstrategie

Die Restriktionskarte enthält eine Vielzahl von Fragmenten, die zur Sequenzierung eingesetzt werden könnten. Welche Restriktionsendonuclease jedoch zur gezielten Fragmentierung der DNA ausgewählt wird, ist davon abhängig, ob das Restriktionsfragment nach seiner Phosphorylierung zur Gewinnung von einseitig markierten Fragmenten entweder einer Strangtrennung unterworfen oder durch ein zweites Enzym gespalten werden kann.

Oft ist es schwierig, eine geeignete Endonuclease zu finden, denn die bei der Spaltung entstehenden Produkte müssen sich in ihrer Größe ausreichend unterscheiden, um in einem Gel trennbar und somit isolierbar zu sein. Sie sollten aber noch groß genug sein, um ein ökonomisches Sequenzieren zu

erlauben. Bei einer regelmäßigen Lesbarkeit des Bandenmusters bis zu zweihundert Nucleotiden ist es sinnvoll, die Sequenz solcher Fragmente zu bestimmen, die mindestens diese Länge erreichen, so daß das Auflösungsvermögen der benutzten Sequenziergele voll ausgeschöpft werden kann. Insbesondere bei kürzeren Restriktionsfragmenten mit etwa dreihundert Basen entsteht durch sekundäre endonucleolytische Spaltung nur ein Bruchstück, das diesen Bedingungen entspricht. Somit ist es notwendig, ein anderes Restriktionsfragment auszuwählen, das den Sequenzbereich des zweiten Spaltproduktes einschließt. Diese Probleme können umgangen werden, wenn sich das entsprechende Restriktionsfragment gelelektrophoretisch in seine Einzelstränge zerlegen und präparativ isolieren läßt. Die Methode der Strangtrennung bietet also strategische Vorteile. Bei kürzeren Stücken kann sogar, wie im Falle des hier verwendeten 180 Basenpaare umfassenden Fragmentes Sau 3A-Sau 3A, auf eine zweite Markierung verzichtet werden, da die Sequenz der komplementären Stränge mit einem Male bestimmt wird. Doch nicht jedes beliebige doppelsträngige Molekül läßt sich mit Hilfe des benutzten Verfahrens separieren. Erst wenn in Vorversuchen der Erfolg einer Strangtrennung nachgewiesen worden ist, kann eine optimale Strategie für die Zerlegung der Gesamt-DNA entwickelt werden. Dabei muß berücksichtigt werden, daß die ersten drei Nucleotide an dem markierten Ende oft nicht eindeutig aus dem Bandenmuster zu bestimmen sind. Zwischen diesen Banden erscheint ein diffuses Hydrazinreaktionsprodukt (39). Um die Sequenz an den betroffenen Schnittstellen festzulegen, werden solche Fragmente für die Nucleotid-

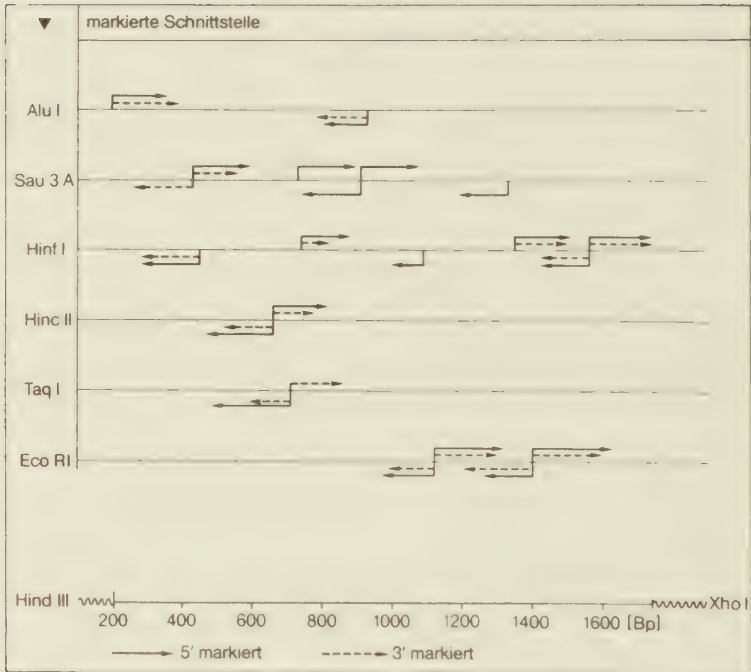


Abbildung 4: Die zur Nucleotidsequenzbestimmung des pho A-Gens benutzten DNA-Fragmente sind durch horizontale Pfeile ober- und unterhalb des Abschnittes Hind III-Xho I dargestellt. Die Querstriche am Ende der Pfeile kennzeichnen die ^{32}P -Markierung. Ihre Länge entspricht der Anzahl der aus den Sequenz-spezifischen Bandenmustern abgelesenen Nucleotide.

analyse ausgesucht, die sich durch überlappende Bereiche an den jeweiligen Enden auszeichnen.

Die zur Sequenzbestimmung des pho A-Gens ausgewählten DNA-Fragmente sind in Abbildung 4 dargestellt. Es wurden auch solche Restriktionsendonuclease-Schnittstellen berücksichtigt, die sich mit dem Fortschreiten der Sequenzanalyse ergeben hatten und sich durch ihre günstige Lage anbieten. Das Enzym Sau 3A beispielsweise wurde bei der Anfertigung der Restriktionskarte nicht benutzt. Durch die in der Abbildung an Position 440, 730, 910 und 1330 gelegenen Schnittstellen entstehen gleich vier Spaltstücke, die sich aufgrund ihrer Überlappungen mit anderen Fragmenten sehr gut zur Sequenzbestimmung eignen, zumal damit fast die Hälfte der Sequenz des pho A-Gens ermittelt werden konnte. Damit entfällt auch die Spaltung mit verschiedenen Enzymen zur Präparation geeigneter Restriktionsfragmente, was natürlich einen geringeren Arbeitsaufwand bedeutet.

3.4 Bestimmung der Nucleotidsequenz des pho A-Gens

Die Bestimmung der Basensequenz des pho A-Gens von E. coli erfolgte durch die "Dimethylsulfat-Hydrazin"-Technik nach Maxam und Gilbert (29) mit Modifikationen von Burton (35) für die Purin-spezifische Reaktion. Die Sequenz wurde sowohl für den codierenden als auch für den nicht-codierenden Strang ermittelt, wobei 3'- und 5'-Markierungen des jeweiligen Restriktionsfragmentes vorgenommen wurden.

Da der Anfang der für die alkalische Phosphatase codierenden Sequenz in dem Abschnitt Hind III-Eco RI des Plasmids pSB 53 lokalisiert worden war, wurde zunächst das größte darin liegende Restriktionsfragment AluI-AluI (730 BP) phosphoryliert und mit Hinc II gespalten. Durch Sequenzanalyse dieser einseitig markierten Spaltprodukte (*Alu I-Hinc II, Hinc II-Alu I*) konnten an beiden Enden etwa einhundertundsechzig Nucleotide bestimmt werden (40). Daß es sich dabei um codierende Regionen des pho A-Gens handelte, wurde durch Translation der Tripletts in ihre zugehörigen Aminosäuren nachgewiesen. Für beide Sequenzen wurden offene Leseraster gefunden, die sich durch das Fehlen von Stopcodonen auszeichnen. Das eine, aus der Basensequenz des *Alu I-Hinc II-Fragmentes resultierend, enthielt sowohl die von Sarthy et al. (41) als Beginn der Präphosphatase identifizierten fünf Aminosäuren Met-Lys-Gln-Ser-Thr als auch die ersten vier Aminosäuren des Strukturproteins, das von Bradshaw et al. (42) vollständig sequenziert worden ist. Das andere, abgeleitet aus der Nucleotidsequenz des Hinc II-Alu I*-Fragmentes, war identisch mit den von Bradshaw et al. (42) gefundenen Aminosäuren 155 bis 193.

Diese Ergebnisse ließen den Schluß zu, daß das pho A-Gen achtzig Nucleotide hinter der in der Restriktionskarte an Position 200 gelegenen Alu I-Schnittstelle mit dem ungewöhnlichen Startcodon GTG beginnt. Das Triplett GUG kann jedoch bei Prokaryonten neben AUG am Anfang einer mRNA für Formylmethionin stehen und somit die Proteinsynthese initiieren (43).

Andererseits war damit auch klar, daß das pho A-Gen innerhalb des Plasmids im Uhrzeigersinn transkribiert wird. Folglich entspricht die Leserichtung nicht der in der Genkarte des E. coli Chromosoms angegebenen, wie von Siewert et al. (18) bereits vermutet worden war (21). Auch Sarthy et al. (45) sowie Wanner et al. (45) hatten diesen Verdacht geäußert.

Die Gesamtsequenz des pho A-Gens und die daraus abgeleitete Aminosäuresequenz ist in Abbildung 5 aufgeführt. Auch die benachbarten Bereiche -"downstream" und "upstream" -, die regulatorische Funktionen übernehmen, sind dargestellt. Die beiden Buchstabenfolgen (A=Adenin, G=Guanin, C=Cytosin, T=Thymin) repräsentieren die komplementären Einzelstränge der DNA. In der Abbildung ist die Sequenz so angeordnet, daß der obere Strang links mit dem 5'-Ende beginnt. Die Numerierung der Nucleotide fängt mit der ersten Base des Startcodons GTG an und erfolgt zum 3'-Ende "downstream" mit positiven, zum 5'-Ende "upstream" hingegen mit negativen Zahlen.

Abbildung 5: Nucleotidsequenz des pro A-Gens sowie seiner angrenzenden Bereiche "upstream" (-80 bis -1) und "downstream" (1417 bis 1464). Die Nucleotide beziehungsweise Basen sind durch ihren jeweiligen Anfangsbuchstaben symbolisiert. Für die Aminosäuresequenz, die sich durch Translation des Nucleotidcodes von Position 1 bis 1416 ergibt, wurde der übliche 3-Buchstaben-Code benutzt.

-80	-70	-60	-50	-40	-30	-20	-10	
5'	CTGTCATAAAGTTGTCTACGGCCGAGACTTATAGTCGCTTTGTTTTTTTAAATGTATTGTACATGGAGAAAAATAAA							
3'	GACAGTATTTCACAGTGCCTGGCTCTGAATATCAGCGAAACAAAATAAAAATTACATAAACATGTACCTCTTTTATT							
1	10	20	30	40	50	60		
GGS AAA CAA AGC ACT ATT GCA CTG GCA CTC TTA CCG TTA CTG TTT ACC CCT GTG ACA AAA								
CAC TTT GTT TCG TGA TAA CGT GAC CGT GAG AAT GCG AAT GAC AAA TGG GGA CAC TGT TTT								
Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr Pro Val Thr Lys								
70	80	90	100	110	120			
GCC CGG ACA CCA GAA ATG CCT GTT CTG GAA AAC GCG GCT GCT CAG GGC GAT ATT ACT GCA								
CGG GCC TGT GGT CTT TAC GGA CAA GAC CTT TTG GCC CGA CGA GTC CCG CTA TAA TGA CGT								
Ala Arg Thr Pro Glu Met Pro Val Leu Glu Asn Arg Ala Ala Gln Gly Asp Ile Thr Ala								
130	140	150	160	170	180			
CCC GGC GGT GCT CGC CGT TTA ACG GGT GAT CAG ACT GCC GCT CTG CGT GAT TCT CTT AGC								
GGG CCG CCA CGA GCG GCA AAT TGC CCA CTA GTC TGA CCG CGA GCA AGA GAA TCG								
Pro Gly Gly Ala Arg Arg Leu Thr Gly Asp Gln Thr Ala Ala Leu Arg Asp Ser Leu Ser								
190	200	210	220	230	240			
GAT AAA CCT GCA AAA AAT ATT ATT TTG CTG ATT GGC GAT GGG ATG GGG GAC TCG GAA ATT								
CTA TTT GGA CGT TTT TTA TAA TAA AAC GAC TAA CCG CTA CCC TAC CCC CTG AGC CTT TAA								
Asp Lys Pro Ala Lys Asn Ile Ile Leu Ile Gly Asp Gly Met Gly Asp Ser Glu Ile								
250	260	270	280	290	300			
ACT GCC GCA CGT AAT TAT GCC GAA GGT GCG GGC GGC TTT TTT AAA GGT ATA GAT GCC TTA								
TGA CCG CGT GCA TTA ATA CCG CTT CCA CCG CCG CCG AAA AAA TTT CCA TAT CTA CGG AAT								
Thr Ala Ala Arg Asn Tyr Ala Glu Gly Ala Gly Gly Phe Phe Lys Gly Ile Asp Ala Leu								
310	320	330	340	350	360			
CCG CTT ACC GGG CAA TAC ACT CAC TAT GCG CTG AAT AAA AAA ACC GGC AAA CCG GAC TAC								
GGC GAA TGG CCC GTT ATG TGA GTG ATA GCG GAC TTA TTT TTT TGG CCG TTT GGC CTG ATG								
Pro Leu Thr Gly Gln Tyr Thr His Tyr Ala Leu Asn Lys Lys Thr Gly Lys Pro Asp Tyr								
370	380	390	400	410	420			
GTC ACC GAC TCG GCT GCA TCA GCA ACC GCG TGG TCA ACC GGT GTC AAA ACC TAT AAC GGC								
CAG TGG CTG AGC CGA CGT AGT CGT TGG CCG ACC AGT TGG CCA CAG TTT TGG ATA TTG CCG								
Val Thr Asp Ser Ala Ala Ser Ala Thr Ala Trp Ser Thr Gly Val Lys Thr Tyr Asn Gly								
430	440	450	460	470	480			
GCG CTG GGC GTC AAT ATT CAC GAA AAA GAT CAC CCA ACG ATT CTG GAA ATG GCA AAA GGC								
CGC GAC CCG CAG CTA TAA GTG CTT TTT CTA GTG CGT TGC TAA GAC CTT TAC CGT TTT CGG								
Ala Leu Gly Val Asp Ile His Glu Lys Asp His Pro Thr Ile Leu Glu Met Ala Lys Ala								

490 500 510 520 530 540
 GCA GGT CTG GCG ACC GGT AAC GTT TCT ACC GCA GAG TTG CAG GAT GCC ACG CCC GCT GCG
 CGT CCA GAC CGC TGG CCA TTG CAA AGA TGG CGT CTG AAC GTC CTA CGG TGC GGG CGA CGC
 Ala Gly Leu Ala Thr Gly Asn Val Ser Thr Ala Glu Leu Gln Asp Ala Thr Pro Ala Ala
 550 560 570 580 590 600
 CTG GTG GCA CAT GTG ACC TCG CGC AAA TGC TAC GGT CCG ACG GCG ACC AGT GAA AAA TGT
 GAC CAC CGT GTA CAC TGG AGC GCG TTT ACG ATG CCA GGC TCG CGC TGG TCA CTT TTT ACA
 Leu Val Ala His Val Thr Ser Arg Lys Cys Tyr Gly Pro Ser Ala Thr Ser Glu Lys Cys
 610 620 630 640 650 660
 CCG GGT AAC GCT CTG GAA AAA GGC GGA AAA GGA TCG ATT ACC GAA CAG CTG CTT AAC GCT
 GGC CCA TTG CGA GAC CTT TTT CCG CCT TTT CCT AGC TAA TGG CTT GTC GAC GAA TTG CGA
 Pro Gly Asn Ala Leu Glu Lys Gly Gly Lys Gly Ser Ile Thr Glu Gln Leu Leu Asn Ala
 670 680 690 700 710 720
 CGT GCC GAC GTT ACG CTT GGC GGC GGC GCA AAA ACC TTT GCT GAA ACG GCA ACC GCT GGT
 GCA CGG CTG CAA TGC GAA CCG CCG CCG CGT TTT TGG AAA CGA CTT TGC CGT TGG CGA CCA
 Arg Ala Asp Val Thr Leu Gly Gly Gly Ala Lys Thr Phe Ala Glu Thr Ala Thr Ala Gly
 730 740 750 760 770 780
 GAA TGG CAG GGA AAA ACG CTG CGT GAA CAG GCA CAG GCG CGT GGT TAT CAG TTG GTG AGC
 CTT ACC GTC CCT TTT TGC GAC GCA CTT GTC CGT GTC GGC GCA CCA ATA GTC AAC CAC TCG
 Glu Trp Gln Gly Lys Thr Leu Arg Glu Gln Ala Gln Ala Arg Gly Tyr Gln Leu Val Ser
 790 800 810 820 830 840
 GAT GCT GCC TCG CTG AAT TCG GTG ACG GAA CCG AAT CAG CAA AAA CCC CTG CTT GGC CTG
 CTA CGA CGG AGC GAC TTA AGC CAC TGC CTT CGC TTA GTC GTT TTT GGG GAC GAA CCG GAC
 Asp Ala Ala Ser Leu Asn Ser Val Thr Glu Ala Asn Gln Gln Lys Pro Leu Leu Gly Leu
 850 860 870 880 890 900
 TTT GCT GAC GGC AAT ATG CCA GTG CGC TGG CTA GGA CCG AAA GCA ACG TAC CAT GGC AAT
 AAA CGA CTG CCG TTA TAC GGT CAC GCG ACC GAT CCT GGC TTT CGT TGC ATT GTA CCG TTA
 Phe Ala Asp Gly Asn Met Pro Val Arg Trp Leu Gly Pro Lys Ala Thr Tyr His Gly Asn
 910 920 930 940 950 960
 ATC GAT AAG CCC GCA GTC ACC TGT ACG CCA AAT CCG CAA CGT AAT GAC AGT GTA CCA ACC
 TAG CTA TTC GGG CGT CAG AGT ACA TGC GGT TTA GGC GTT CGA TTA CTG TCA CAT GGT TGG
 Ile Asp Lys Pro Ala Val Thr Cys Thr Pro Asn Pro Gln Arg Asn Asp Ser Val Pro Thr
 970 980 990 1000 1010 1020
 CTG GCG CAG ATG ACC GAC AAA GCC ATT GAA TTG TTG AGT AAA AAT GAG AAA GGC TTT TTC
 GAC CGC GTC TAC TGG CTG TTT CGG TAA CTT AAC AAC TCA TTT TTA CTC TTT CCG AAA AAG
 Leu Ala Gln Met Thr Asp Lys Ala Ile Glu Leu Leu Ser Lys Asn Glu Lys Gly Phe Phe
 1030 1040 1050 1060 1070 1080
 CTG CAA GTT GAA GGT GCG TCA ATC GAT AAA CAG GAT CAT GCT GCG AAT CCT TGT GGG CAA
 GAC GTT CAA CTT CCA CGC AGT CTA TTT GTC CTA TTA GCA CGC TTA GAT CGA ACA CCC GTT
 Leu Gln Val Glu Gly Ala Ser Ile Asp Lys Gln Asp His Ala Ala Asn Pro Cys Gly Gln
 1090 1100 1110 1120 1130 1140
 ATT GGC GAG ACG GTC GAT CTC GAT GAA GCC GTA CAA CGG GCG CTG GAA TTC GCT AAA AAG
 TAA CCG CTC TGC CAG CTA GAT CTT CGG CAT GTT GCC GAC CTT AAG CGA TTT TTC
 Ile Gly Glu Thr Val Asp Leu Asp Glu Ala Val Gln Arg Ala Leu Glu Phe Ala Lys Lys
 1150 1160 1170 1180 1190 1200
 GAG GGT AAC ACG CTG GTC ATA GTC ACC GCT GAT CAC GCC CAC GCC AGC CAG ATT GTT GCG
 CTC CCA TTG TGC GAC CAG TAT CAG TGG CGA CTA GTG CGG GTG CGG TCG GTC TAA CAA CGC
 Glu Gly Asn Thr Leu Val Ile Val Thr Ala Asp His Ala His Ala Ser Gln Ile Val Ala

1210	1220	1230	1240	1250	1260
CCG GAT ACC AAA GCT CCG GGC CTC ACC CAG GCG CTA AAT ACC AAA GAT GGC GCA GTG ATG					
GGC CTA TGG TTT CGA GGC GGC GAG TGG GTC CCG GAT TTA TGG TTT CTA CCG CGT CAC TAC					
Pro Asp Thr Lys Ala	Pro Gly Leu Thr	Gln Ala Leu Asn Thr	Lys Asp Gly Ala	Val Met	
1270	1280	1290	1300	1310	1320
GTG ATG AGT TAC GGG ACC TCC GAA GAG GAT TCA CAA GAA CAT ACC GGC AGT CAG TTG CGT					
CAC TAC TCA ATG CCC TGG AGG CTT CTC CTA AGT GTT CTT GTA TGG CCG TCA GTC AAC GCA					
Val Met Ser Tyr Gly	Asn Ser Glu Glu	Asp Ser Gln Glu His Thr	Gly Ser Gln Leu	Arg	
1330	1340	1350	1360	1370	1380
ATT GCG GCG TAT GGC CCG CAT GGC GCC AAT GTT GTT GGA CTG ACC GAC CAG ACC GAT CTC					
TAA CGC CGC ATA CCG GGC GTA CGG CGG TTA CAA CAA CCT GAC TGG CTG GTC TGG CTA GAG					
Ile Ala Ala Tyr Gly	Pro His Ala Ala	Asn Val Val Gly	Leu Thr Asp Gln Thr	Asp Leu	
1390	1400	1410	1420	1430	1440
TTC TAC ACC ATG AAA GCC GCT CTG GGG CTG AAA TAA AACCGCGCGCGGCACTGAATTTTGGCGCGG					
AAG ATG TGG TAC TTT CGG CGA GAC CCC GAC TTT ATT TTGGCGCGCGCGCTCACTTAAAAACGCGCGG					

Phe Tyr Thr Met Lys Ala Ala Leu Gly Leu Lys Stop

1450 1460

GTGGTTTTTGCGTTAGCA

CACCAAAAACGCAATCGT

3.4.1 Schwierigkeiten bei der Interpretation des Banden- musters

Bei der Auswertung des autoradiographischen Bandenmusters zur Bestimmung der Sequenz wurden mitunter Artefakte beobachtet, die keine eindeutige Zuordnung der Nucleotide erlaubten.

Normalerweise nimmt der Abstand zwischen den Banden von unten nach oben gleichmäßig ab. Sehr häufig konnten jedoch Unregelmäßigkeiten in der Bandenabfolge in der Form festgestellt werden, daß erwartete Banden fehlten oder die Bandendistanz innerhalb einer Gelspur oder in den vier parallelen Gelbahnen zu groß oder zu klein war. Diese von Brown und Smith (46) als Kompression beschriebenen Effekte beruhen auf stabilen Paarungen der 3'-endständigen Basen mit einem komplementären Bereich im selben Einzelstrang. In Abbildung 6 ist eine solche Struktur neben dem eindeutig zu interpretierenden Bandenmuster des komplementären Stranges dargestellt. Die Sequenzbestimmung des Komplementärstranges ist nämlich die einfachste Methode, solche kritischen Bereiche innerhalb eines Restriktionsfragmentes aufzuklären. Eine andere Möglichkeit besteht in der Auswahl eines Fragmentes, das die betreffende Sequenz an seinem markierten Ende hat, so daß diese in den Bereich höherer Auflösung gerückt wird. Auch durch eine Erhöhung der Harnstoffkonzentration und der Temperatur während der Elektrophorese können diese Faltungen innerhalb des Moleküls vermieden werden. Eine höhere Harnstoffkonzentration als die eingesetzte würde jedoch die Löslichkeitsgrenzen überschreiten. Maxam und Gilbert (29) empfehlen eine

Temperatur von mindestens 50°C, was durch eine ausreichend hohe Spannung während der Elektrophorese gewährleistet werden kann. Eine stärkere Aufheizung ist jedoch durch die Thermostabilität der Glasplatten, die das Gelbett formen, begrenzt.

6a

6b

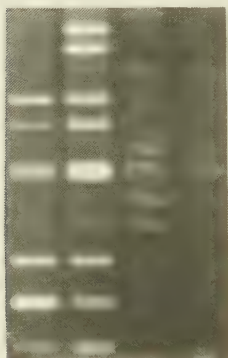


Abbildung 6: Das Fragment Hind III-Taq I wurde mit ^{32}P sowohl an seinem 5'- als auch an seinem 3'-Ende markiert und mit Hae II gespalten. Die Spaltprodukte wurden anschließend durch Gelelektrophorese getrennt. Das einseitig 3' beziehungsweise 5' markierte Fragment Hae II-Taq I* wurde aus der Gelmatrix isoliert und einer Sequenzanalyse unterzogen. Die Abbildung zeigt Ausschnitte der zugehörigen Sequenzspezifischen Bandenmuster (20%iges Polyacrylamidgel). Die Sequenz des 3' markierten Fragmentes (6a) läßt sich aus der Autoradiographie aufgrund der zwischen den Cytosin- und Guaninbasen erfolgten Paarungen (markierte Region I) nicht eindeutig bestimmen. Das entsprechende Bandenmuster des 5' markierten Fragmentes (6b) zeigt in dem betreffenden Bereich (II) keine Kompressionseffekte, so daß die Basenfolge in der richtigen Reihenfolge abgelesen werden kann.

Weitere Schwierigkeiten bei der Interpretation des Bandenmusters entstanden durch unspezifische Banden in allen Gelbahnen. Diese lassen sich auf Kontaminationen mit anderer markierter DNA oder auf nicht absolute Basenspezifität der Hydrolysebedingungen zurückführen (29). Sie stellen sich jedoch meist als so schwach ausgeprägte Banden dar, daß aufgrund des Intensitätsunterschiedes die Auswertung mit ausreichender Sicherheit durchgeführt werden kann.

Verunreinigungen durch Salze machten sich durch nach unten verjüngende Gelspuren bemerkbar und bewirkten eine Beeinträchtigung der Lesbarkeit der betroffenen Banden mit zunehmender Entfernung vom ^{32}P markierten Ende. Der Versuch mußte dann wiederholt werden.

In der Guanin-spezifischen Spur fehlten häufig die Banden im oberen Bereich des Gels. Erst eine Reduzierung sowohl der vorgeschriebenen Reagenzmenge als auch der Reaktionszeit führte zu einer geringeren Spaltungsrate, so daß mehr Nucleotide am Bandenmuster abgelesen werden konnten. Dabei wurde beobachtet, daß sich bei mehreren direkt aufeinanderfolgenden Desoxy-Guanylresten in der DNA die letzten beiden Basen oder nur die letzte Base dieses Blocks in der Autoradiographie schwächer darstellten.

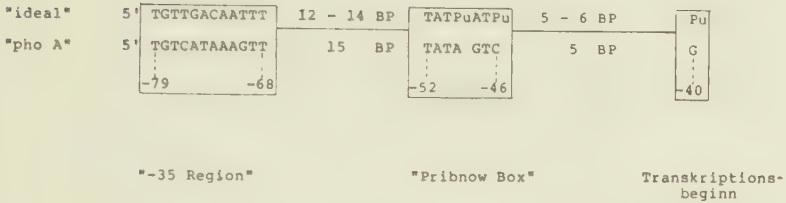
3.4.2 Promotor

Die Transkription eines Gens in *E. coli* wird katalysiert durch die DNA-abhängige RNA-Polymerase. Dieses Enzym leitet den mRNA-Syntheseprozess jedoch nur ein, wenn dem zu kopierenden DNA-Molekül eine besondere Nucleotidsequenz vorangeht (47). Eine Vielzahl solcher als Promotoren bezeichneten Initiationsstellen sind in den letzten Jahren untersucht worden (48). Dabei wurde festgestellt, daß alle diese Sequenzen zwei gleiche oder doch sehr ähnliche Abschnitte enthalten.

Pribnow (49) sowie Schaller et al. (50) fanden eine sieben Basenpaare umfassende Sequenz, mit ihrem Zentrum etwa zehn Nucleotide "upstream" vor dem Startpunkt der Transkription liegend, die bei Prokaryonten fast ausnahmslos mit einer Purinbase beginnt (51). 35 Nucleotidpaare von dieser Stelle entfernt wurde eine Sequenzfolge aus 12 Basen nachgewiesen (52). Es wurde postuliert, daß diese sogenannte "-35 Region" als Erkennungsstelle für die RNA-Polymerase dient, während die "Pribnow Box" als Bindungsstelle des Enzyms gilt (53).

Anhand dieser Homologien wurde eine Promotor-Modellsequenz erstellt (53). Der DNA-Abschnitt von Position -80 bis -1, der dem *pho A*-Gen vorgeschaltet ist, wurde auf Übereinstimmung mit dieser Mustersequenz untersucht. Die daraus resultierenden ähnlichen Sequenzfolgen sind zusammen mit den idealisierten Formen für die Erkennungs- und Bindungsstelle aufgezeichnet. Ihre Lage innerhalb der ermittelten Sequenz ist aus den Positionsnummern ersichtlich; es ist, der Konvention entspre-

chend, jeweils der nonsense-Strang der DNA wieder-
gegeben.



3.4.3 Bindungsstelle für Ribosomen

Natürliche mRNA-Moleküle von Prokaryonten besitzen zwischen ihrem 5'-Ende und ihrem Startcodon eine nicht translatierte Sequenz (54). Etwa zehn Nucleotide vor dem Startcodon wurde stets eine der Sequenz 5' GGAGGU 3' entsprechende Nucleotidfolge gefunden (55). Shine und Dalgarno (56) stellten fest, daß das 3'-Ende der 16S ribosomalen RNA (GAUCCUCCUUA_{OH}3') eine komplementäre Basenfolge enthält. Daraus schlossen sie, daß diese kurze Nucleotidsequenz der mRNA die Bindungsstelle für Ribosomen darstellt und somit zum Starten der Proteinbiosynthese erforderlich ist.

Dem pho A-Gen sollte ebenfalls ein solcher nicht für das Protein codierender Abschnitt vorgeschaltet sein. Unter der Voraussetzung, daß die Transkription der DNA fünf Nucleotide hinter der "Pribnow Box" mit dem Purin Guanin beginnt, dehnt sich dieser Bereich über vierzig Basen bis zum Startcodon der Proteinbiosynthese GTG aus. Bei der

Suche nach einer solchen Purin-reichen Nucleotidsequenz in entsprechender Entfernung zum Startcodon wurde eine Folge von vier Basen gefunden, die mit der angegebenen Sequenz des Transkripts übereinstimmen. Die Sequenz 5' GGAG 3' reicht von der Position -12 bis -9 und kann mit großer Wahrscheinlichkeit als die Ribosomenbindungsstelle des pho A-messengers angesehen werden.

3.4.4 Terminator

Der Mechanismus, der zum Abbruch der mRNA-Synthese führt, ist weniger aufgeklärt als die Initiationsreaktion. Die Frage, ob eine spezifische Nucleotidsequenz am Ende eines zu transkribierenden DNA-Abschnittes als Terminationssignal ausreicht oder ob weitere regulatorische Proteine hierzu notwendig sind, konnte bisher nicht eindeutig entschieden werden.

In vielen untersuchten Terminatoren wurden Bereiche mit hohem G-C-Anteil gefunden, denen sich eine Serie von vier bis acht Desoxy-Thymidylresten anschließt. Innerhalb dieser Folge wird die Transkription beendet (48). Auch der sich dem pho A-Gen anschließende Abschnitt enthält eine G-C-reiche Sequenz, der eine Reihe von fünf Thymidinbasen folgt. Somit könnte dieser Bereich von Position 1440 bis 1455 für die Beendigung der Transkription des pho A-Gens sorgen.

Mögliche Terminatorsequenz:

5'	CGCGCCGGTGGT	3'
	1440	1455

Wenn damit auch ein Mitwirken eines für die Termination notwendigen Regulationsproteins nicht ausgeschlossen werden kann, so scheint doch zumindest der Terminationsfaktor Rho an diesem Prozeß nicht beteiligt zu sein, denn bei den obligatorisch Rho-Faktor abhängigen Terminatoren wurden solche Folgen von Thymidinbasen nicht festgestellt (48). In allen bisher analysierten Terminatoren wurden jedoch invertiert wiederholte Sequenzabschnitte gefunden, die durch intrasträngige Basenpaarung komplementärer Nucleotide haarnadel-förmige Sekundärstrukturen einnehmen können (48). Ob diese Palindrome ebenfalls Kennzeichen des für das pho A-Gen postulierten Terminators sind, sollte sich bei der systematischen Suche nach solchen Strukturelementen zeigen (siehe Kapitel 3.6.2).

3.5 Aminosäuresequenz

Die aus dem Nucleotidcode des pho A-Gens abgeleitete Aminosäuresequenz kann Abbildung 5 entnommen werden. Die Aminosäurezusammensetzung dieses Genproduktes, das aus 471 Bausteinen besteht, ist in Tabelle 2 angegeben. Daraus ergibt sich ein Molekulargewicht von 49404. Bradshaw et al. (42) hingegen ermittelten bei der direkten Bestimmung der Aminosäuresequenz der alkalischen Phosphatase 449 Aminosäuren mit einem Molekulargewicht von 47029. Das Translationsprodukt ist demnach um zweiundzwanzig Aminosäuren länger als das Protein selbst, das aus dem periplasmatischen Raum isoliert worden ist. Dieser Befund wird bestätigt

durch in vitro und in vivo Versuche zur Synthese der alkalischen Phosphatase von *E. coli* (22, 57). Inouye und Beckwith (22) postulierten, daß auch dieses Enzym zunächst als größere Vorstufe synthetisiert wird, wie das für eukaryontische, aber auch prokaryontische Exoproteine bereits nachgewiesen werden konnte (58, 59). Diese sogenannten Präproteine bestehen aus dem eigentlichen Protein und einem aminoterminalen Signalpeptid von etwa fünfzehn bis dreißig überwiegend hydrophoben Aminosäuren, das während oder nach der Sekretion enzymatisch abgespalten wird.

Sarthy et al. (41) identifizierten bei der Untersuchung eines Hybridproteins die ersten fünf Aminosäuren des Signalpeptids der alkalischen Phosphatase. Sie stehen in Übereinstimmung mit der aus der ermittelten DNA-Sequenz abgeleiteten Aminosäuresequenz Met-Lys-Gln-Ser-Thr. Zudem ist dieser aminoterminaler Abschnitt, der nicht Bestandteil der durch Analyse des Proteins ermittelten Aminosäuresequenz ist, reich an hydrophoben Aminosäuren, so daß die Aminosäuren 1 bis 21 als Signalpeptid der alkalischen Phosphatase angesehen werden können.

Die Aminosäure Arginin an Position 22 repräsentiert die erste Aminosäure des Isozyms I. Die alkalische Phosphatase kommt nämlich in drei verschiedenen Formen vor, die sich elektrophoretisch und chromatographisch auftrennen lassen (60). Isozym III unterscheidet sich nur durch das Fehlen dieses Argininrestes vom Isozym I, während Isozym II ein Heterodimer der beiden Polypeptidketten darstellt (61).

Tabelle 2: Aminosäurezusammensetzung der alkali-
schen Phosphatase

Klasse	Aminosäure	Anzahl im Protein
unpolar	Alanin	67
	Isoleucin	17
	Leucin	44
	Methionin	8
	Phenylalanin	9
	Prolin	23
	Valin	25
	Tryptophan	3
polar	Asparagin	20
	Cystein	4
	Glutamin	23
	Glycin	45
	Serin	23
	Tyrosin	11
	Threonin	43
sauer	Asparaginsäure	28
	Glutaminsäure	24
basisch	Arginin	14
	Histidin	10
	Lysin	30

3.6 Analyse des pho A-Gens

3.6.1 Erkennungssequenzen für Restriktionsendonucleasen

Die in Abbildung 5 angegebene Nucleotidsequenz wurde von Position -80 bis 1416 systematisch durchsucht nach dem Vorkommen von Erkennungssequenzen für Restriktionsenzyme. Für die nachfolgend tabellierten Endonucleasen wurden solche spezifischen Nucleotidkombinationen gefunden. Es sind sowohl ihre Positionen als auch in Klammern ihre Abstände voneinander beziehungsweise bis zum Terminationscodon aufgelistet, wobei die Positionsnummer jeweils die erste Base der Erkennungssequenz angibt.

Für die folgenden zur Untersuchung eingesetzten Enzyme konnten keine passenden Nucleotidsequenzen ermittelt werden:

Acc I, Asu II, Ava I, Ava III, Ava E, Bal I, Bam HI, Bga I, Bgl I, Bgl II, Ecl II, Eco B, Eco K, Eco RV, Gdi I, Hgi CI, Hgi EII, Hind III, Hpa I, Kpn I, Mst I, Nae I, Nar I, Nru I, Pst I, Pvu I, Rru I, Rsh I, Sac I, Sac II, Sal I, Sau I, Sma I, Sna I, Uba I, Xba I, Xho I, Xho II, Xma I, Xmn I.

Ein Teil dieser Nucleasen wurde zur vollständigen Spaltung der Fragmente Hind III-Eco RI und Eco RI-Xho I eingesetzt. Auch dabei wurden keine Schnittstellen für die betreffenden Enzyme festgestellt (siehe Tabellen 1a und 1b).

Tabelle 3: Positionen und Abstände bekannter Erkennungssequenzen für Restriktionsendonucleasen

Acy I	427	(990)				
Alu I	647	(566)	1212	(205)		
Asu I	574	(301)	874	(345)	1218	(116)
	1333	(84)				
Atu I	389	(572)	960	(268)	1227	(190)
Ava II	574	(301)	874	(543)		
Bbv I	97	(277)	373	(163)	535	(114)
	648	(91)	738	(47)	784	(277)
	1060	(357)				
Bcl I	147	(1024)	1170	(247)		
Cla I	633	(269)	901	(142)	1042	(375)
Dde I	101	(75)	175	(1242)		
Eco RI	795	(332)	1126	(291)		
Eco RI'	795	(136)	930	(136)	1065	(62)
	1126	(162)	1287	(130)		
Eco RI*	237	(17)	253	(544)	796	(194)
	989	(92)	1080	(48)	1127	(290)
Eco RII	389	(572)	960	(268)	1227	(190)
Fnu H	97	(61)	157	(88)	244	(29)
	272	(102)	373	(106)	478	(58)
	535	(114)	648	(33)	680	(4)
	683	(56)	738	(47)	784	(277)
	1060	(265)	1324	(19)	1342	(55)
	1396	(21)				
Gdi II	-63	(1479)				
Hae I	834	(583)				
Hae II	1119	(112)	1230	(187)		
Hae III	-62	(897)	835	(385)	1219	(115)
	1333	(84)				
Hga I	428	(609)	1036	(381)		

Hgi AI	128	(1289)				
Hha I	328	(92)	419	(3)	421	(118)
	538	(24)	561	(21)	581	(106)
	686	(72)	757	(108)	864	(101)
	964	(157)	1120	(79)	1198	(34)
	1231	(20)	1250	(167)		
Hind II	393	(1024)				
Hinf I	169	(61)	229	(139)	367	(93)
	459	(355)	813	(253)	1065	(224)
	1288	(129)				
Hpa II	63	(31)	93	(30)	122	(187)
	308	(37)	344	(9)	352	(47)
	398	(97)	494	(108)	601	(601)
	1201	(16)	1216	(89)	1304	(113)
Hph I	145	(218)	360	(89)	450	(269)
	718	(57)	774	(28)	801	(117)
	917	(247)	1163	(61)	1223	(38)
	1260	(157)				
Mbo II	1282	(98)	1379	(38)		
Mnl I	557	(232)	788	(434)	1221	(196)
Nci I	62	(32)	93	(29)	121	(188)
	308	(294)	601	(616)	1216	(201)
Nco I	891	(526)				
Pvu II	646	(771)				
Rsa I	-18	(906)	888	(36)	923	(30)
	952	(160)	1111	(306)		
Sau 3A	148	(301)	448	(185)	632	(423)
	1054	(43)	1096	(76)	1171	(205)
	1375	(42)				
Sfa NI	292	(85)	376	(148)	523	(259)
	781	(636)				

Sph I	1338	(79)				
Taq I	431	(204)	634	(269)	902	(142)
	1043	(52)	1094	(7)	1100	(317)
Tha I	420	(141)	560	(23)	582	(177)
	758	(659)				
Tth 3I	355	(1062)				
Tth 3II	830	(10)	839	(578)		
Xma III	-63	(1479)				

3.6.2 Basenverteilung

Bei der Ermittlung der Basenzusammensetzung des pho A-Gens wurde ein größerer Anteil an G-C als an A-T Nucleotiden gefunden

$$A = T = 659 \text{ Basen} = 46,5\%$$

$$G = C = 757 \text{ Basen} = 53,5\%$$

Mehr Information über den Aufbau des Gens als die Gesamtzusammensetzung gibt die Basenverteilung in kleinen Intervallen. Zu diesem Zweck wurde der relative Gehalt an A-T und G-C Nucleotiden in 59 Intervallen zu je 24 Basenpaaren ermittelt, mit dem ersten Nucleotid des Startcodons beginnend. Der Anfang und die Größe dieser Intervalle sind willkürlich festgesetzt worden. In gleicher Weise wurden auch die benachbarten Bereiche untersucht, die als Promotor und Ribosomenbindungsstelle sowie

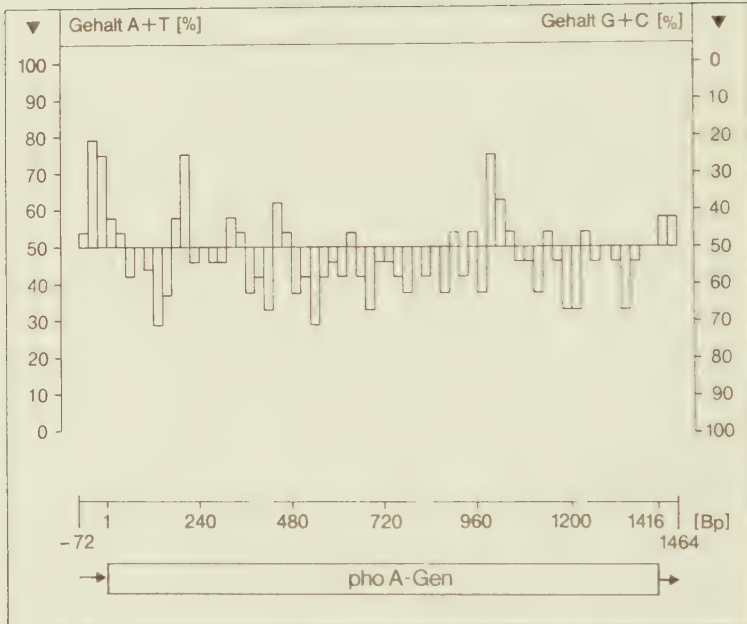


Abbildung 7: Verteilung des relativen Gehaltes an A-T und G-C Nucleotidpaaren im pho A-Gen sowie seinen angrenzenden Bereichen "upstream" und "downstream" (durch Pfeile markiert)

als Terminator beschrieben worden sind. Dabei konnten die ersten Basen des Promotors von Position -80 bis -73 aufgrund der Intervalleinteilung nicht berücksichtigt werden.

In Abbildung 7 läßt sich eine ungleiche Verteilung von G-C gegenüber A-T Nucleotiden im pho A-Gen deutlich erkennen. In der dem Strukturgen vorgeschalteten Region (-72 bis -1) überwiegen A-T Paare (68%), während im Strukturgen eher G-C Basenpaarungen gehäuft vorkommen.

3.6.3 Häufigkeit des Auftretens von Codewörtern

Der genetische Code ist "degeneriert", d.h. für die Codierung der meisten Aminosäuren steht mehr als ein Triplet zur Verfügung. Bei der Analyse verschiedener mRNA- und DNA-Moleküle wurde festgestellt, daß die Zelle von dem Angebot synonymen Codewörter selektiv Gebrauch macht (62, 63). In der Absicht, ähnliche Informationen für das pho A-Gen zu bekommen, wurde die Häufigkeit des Auftretens der verwendeten Triplets ermittelt. Wie aus Tabelle 4 zu ersehen ist, treten mit Ausnahme der Triplets AGA, AGG und CGA, die die Aminosäure Arginin determinieren, sowie der beiden Stopcodone TAG und TGA alle Codewörter mindestens einmal auf. Doch auch hier läßt sich eine ungleichmäßige Ausnutzung der einzelnen Triplets registrieren.

Tabelle 4: Häufigkeit des Auftretens von Codewörtern im pho A-Gen

Phe	TTT	6	Ser	TCT	2	Tyr	TAT	5	Cys	TGT	3
	TTC	3		TCC	1		TAC	6		TGC	1
Leu	TTA	4		TCA	4	och	TAA	1	umb	TGA	0
	TTG	6		TCG	6	amb	TAG	0	Trp	TGG	3

Leu	CTT	5	Pro	CCT	4	His	CAT	5	Arg	CGT	8
	CTC	4		CCC	4		CAC	5		CGC	3
	CTA	3		CCA	5	Gln	CAA	8		CGA	0
	CTG	23		CCG	10		CAG	15		CGG	3

Ile	ATT	13	Thr	ACT	5	Asn	AMT	13	Ser	AGT	5
	ATC	2		ACC	25		AAC	7		AGC	5
	ATA	2		ACA	2	Lys	AAA	28	Arg	AGA	0
Met	ATG	8		ACG	11		AAG	2		AGG	0

Val	GTT	7	Ala	GCT	17	Asp	GAT	20	Gly	GGT	13
	GTC	7		GCC	17		GAC	8		GGC	21
	GTA	2		GCA	17	Glu	GAA	19		GGA	5
	GTG	9		GCG	16		GAG	5		GGG	6

Für die RNA des Virus MS 2 sowie für die Gene F und G des Bakteriophagen Φ X 174 konnte nachgewiesen werden, daß Codone, deren dritte Base ein U ist, dominieren (64,65). Die Auswahl der Base an dieser Position ist nach der Hypothese von Grantham et al. (66), die aus einem Vergleich von 29 bakteriellen mRNA-Sequenzen resultiert, abhängig von der Paarungsenergie zwischen Codon und Anticodon. Diese sollte bei stark exprimierten Genen einen geringeren Wert erreichen als bei schwach exprimierten, denn eine niedrige Bindungsstärke erlaubt einen schnellen Wechsel und folglich eine hohe Syntheserate. Wie Grantham et al. feststellten, wird dieses Prinzip bei dem Einbau der Pyri-

midine Cytosin und Uracil als dritter Base enthalten, nicht aber bei den Purinbasen Adenin und Guanin. Vielmehr wird das Triplet CCG in stark exprimierten Genen bei der Codierung der Aminosäure Prolin favorisiert, obwohl es die höchste Paarungsenergie aufweist.

Dieses Codon wird im pho A-Gen ebenfalls häufiger verwendet als das entsprechende mit der Purinbase Adenin endende Triplet. Ebenso treffen auch Grantham's Beobachtungen, daß die Triplets AAA für Lysin, GAA für Glutaminsäure, CAG für Glutamin und CTG für Leucin in stark exprimierten Genen bevorzugt werden, auf das pho A-Gen zu. Daraus könnte gefolgert werden, daß dieses DNA-Molekül beziehungsweise sein Transkript einer hohen Syntheserate unterliegt. Eine Analyse der Triplets unter Berücksichtigung der Pyrimidinbasen an dritter Position läßt eine solche Festlegung jedoch nicht zu. Ein Vergleich der T/C-Verhältnisse mit den von Grantham et al. angegebenen Werten läßt das Gen mal in der Gruppe der schwach exprimierten, mal in der Gruppe der stark exprimierten Moleküle erscheinen. Wie aus der Tabelle zu entnehmen ist, bestehen auch bei den Codewörtern für die Aminosäuren Valin, Alanin, Serin, Histidin und Prolin keine Präferenzen bezüglich der Pyrimidinbasen. Somit ist die Ursache für die selektive Benutzung mancher Codone im pho A-Gen unter den genannten Aspekten nicht eindeutig zu klären. Vielleicht hat die offensichtlich nicht zufällige Verteilung der verschiedenen degenerierten Triplets eher eine physiologische Bedeutung. So postuliert Ikemura (67), daß die Codewörter in *E. coli* häufiger ausgewählt werden, deren tRNA-

Moleküle mit dem passenden Anticodon in hoher Konzentration in der Zelle vorliegen.

3.6.4 Invertiert repetierte Sequenzen

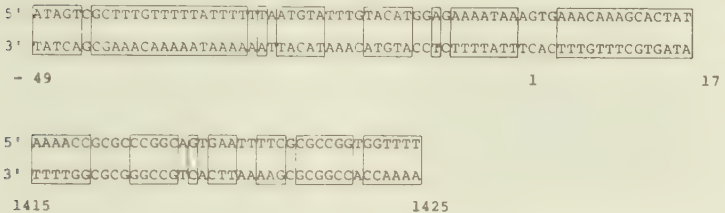
Inverse Repetitionen sind Nucleotidsequenzen, die zueinander komplementär und gegenläufig sind. Das bedeutet, daß sich die Sequenz des oberen Stranges von links nach rechts wie die des unteren Stranges von rechts nach links liest. Demzufolge können sich diese Basenfolgen aneinanderlagern, so daß charakteristische Sekundärstrukturen mit einem doppelsträngigen mittleren Stamm und zwei einzelsträngigen Schleifen entstehen (englisch hairpin).

Das Auftreten solcher symmetrischen Abschnitte in Kontrollregionen des Bakteriengenoms ließ vermuten, daß diese Sequenzen durch ihre exponierte Stellung innerhalb des linearen Doppelstranges mit Regulatorproteinen interagieren (68). Gierer (69) schlug in seinem Modell vor, daß durch die Bindung der Proteine die kreuzförmige Struktur stabilisiert wird, so daß im Falle von Repressoren die Passage der RNA-Polymerase blockiert ist. Wie bereits erwähnt, sind solche Palindrome für die Termination der mRNA-Synthese obligatorisch. Sie sind jedoch nicht zur Kennzeichnung von Promotorregionen erforderlich, da diese selbstkomplementären Nucleotidfolgen nicht in allen bisher bekannten Promotoren gefunden wurden.

Die gesamte in Abbildung 5 angegebene Nucleotidsequenz wurde systematisch nach invertiert repetierten Sequenzen durchsucht. Dabei zeigten sich größere Sequenzblöcke in den Bereichen "upstream"

und "downstream" vom pho A-Gen, also in den Regionen, die der Regulation der Genexpression dienen könnten. Abbildung 8 zeigt die Nucleotidsequenzen für beide Abschnitte, die durch ihre Positionsnummern gekennzeichnet sind. Die Palindrome sind eingerahmt.

Abbildung 8: Invertiert repetierte Sequenzen



4. Diskussion

4.1 Restriktionskarte

Für die Fragmente Hind III-Eco RI und Eco RI-Xho I wurden die Schnittstellen mit verschiedenen Restriktionsendonucleasen experimentell bestimmt und entlang dem DNA-Molekül angeordnet (Abbildung 2a und 2b). Ein Vergleich mit den aus der Nucleotidsequenz des pho A-Gens zu postulierenden Spaltstellen zeigt für die Enzyme Taq I, Hha I und Hpa II keine Übereinstimmung in Anzahl und Größe der Spaltprodukte.

Während das nach vollständiger Spaltung und elektrophoretischer Auftrennung gefundene Bandenmuster nur eine Schnittstelle für Taq I innerhalb des Fragmentes Hind III-Eco RI erwarten läßt, ergibt sich aus der Basensequenz eine weitere potentielle Schnittstelle an Position 634. Diese Diskrepanz ist nicht auf eine fehlerhafte Bestimmung der Nucleotidsequenz an dieser Stelle zurückzuführen, denn das zugehörige sequenzspezifische Bandenmuster ist eindeutig zu interpretieren, wie Abbildung 9 zeigt. Als Erklärung kommt eine Methylierung der Erkennungssequenz 5' TCGA in Frage, da diese Modifikation zur Resistenz gegenüber Restriktionsendonucleasen führt (70). Die prokaryontische Zelle hat mit den korrespondierenden Modifikationsmethylasen, d.h. zu jeder Restriktionsendonuclease gehört ein Modifikationsenzym, die beide mit derselben Nucleotidfolge interagieren, einen wirksamen Mechanismus entwickelt, um den Abbau ihres Genoms durch endogene Restriktionsenzyme zu verhindern. Die Methylierung erfolgt entweder am Stickstoffatom 6 des Adenin oder am



Abbildung 9:

Die Sequenzen der komplementären Stränge des von Position 393 bis 647 reichenden DNA-Fragmentes Hinc II-Alu I* wurden bestimmt. Die zugehörigen autoradiographischen Bandenmuster (20%iges Polyacrylamidgel) zeigen die Erkennungssequenz 5'TCGA (I) beziehungsweise 3'AGCT (II) für das Restriktionsenzym Taq I, deren Adeninreste wahrscheinlich methyliert sind (durch einen Stern markiert). Diese Modifikation führt zur Resistenz dieser Sequenz gegenüber der Restriktionsendonuclease.

Kohlenstoffatom 5 des Cytosin unter Verwendung von S-Adenosyl-Methionin (71). Sato et al. (72) stellten fest, daß in der Sequenz 5'TCGA der Adeninrest durch die Tag I-Methylase zu 6-Methyladenin umgesetzt werden kann. Wahrscheinlich liegt diese Base in beiden Strängen der genannten Position in ihrer methylierten Form vor.

Aufgrund der Sequenz ergeben sich in demselben Fragment für Hpa II neun und für HhaI elf Schnittstellen. Daraus resultiert jeweils ein Restriktionsfragment mehr als das bei vollständiger Verdauung entstandene Bandenmuster aufweist (siehe Tabelle 1a). Ein solcher Befund könnte ebenfalls durch eine Methylierung hervorgerufen werden. Die Erkennungssequenz des Enzyms Hha I 5'GCGC bietet die Möglichkeit zweier methylierter Cytosinreste; in der Erkennungssequenz 5'CCGG für Hpa II hingegen kann nur das innen liegende Cytosin zu Desoxy 5-Methylcytosin modifiziert sein (73, 74). Da 5-Methylcytosin zu langsam mit Hydrazin reagiert, wird die DNA bei der Nucleotidsequenzbestimmung nach Maxam und Gilbert (29) an dieser Stelle kaum gespalten, so daß die Gegenwart eines solchen Restes eine Lücke in dem sequenzspezifischen Bandenmuster hinterläßt (39). Fehlende Banden konnten jedoch an allen möglichen Restriktionsendonuclease-Schnittstellen nicht gefunden werden. Unter Umständen könnte hierfür eine unvollständige Methylierung verantwortlich sein, die aber schon zum Schutz vor Endonuclease-Aktivität ausreichen könnte (75). Somit ist die Frage, ob eine Methylierung einzelner Hha I- und Hpa II-Erkennungssequenzen zu unterschiedlichen Ergebnissen bezüglich der Anzahl der Spaltstellen führt, nicht abschließend zu beantworten.

4.2 Promotor

In allen prokaryontischen Promotoren wurden zwei homologe Bereiche gefunden, die sogenannte "-35 Region" sowie die "Pribnow oder TATA Box", etwa 35 beziehungsweise 10 Nucleotide vor dem Startpunkt der Transkription liegend (53). Beide Abschnitte wurden in der Basensequenz, die dem *pho A*-Gen vorangeht, identifiziert. Zwei Nucleotide der "TATA Box" (-46 bis -52) sind nicht mit der von Pribnow geforderten Sequenz identisch. Ihr Abstand zum Initiationscodon der Proteinbiosynthese (45 BP) ist vergleichbar mit dem im *lac*-Operon (43 BP) (76). Die 12 Nucleotide umfassende Sequenz der "-35 Region" (-68 bis -79) weicht in fünf Basen von der Idealsequenz ab. Unter der Voraussetzung, daß diese Übereinstimmungen mit der Modellsequenz bezüglich der Basenabfolge und -anzahl ausreichen, um die RNA-Polymerase mit hoher Affinität zu binden und damit die RNA-Synthese einzuleiten, kann die von Position -39 bis -79 reichende Sequenz als Promotor des *pho A*-Gens definiert werden. Dann stellt die Base Guanin mit der Positionsnummer -40 den Startpunkt der Transkription dar, denn die RNA-Synthese beginnt bei Prokaryonten fast ausnahmslos mit einer Purinbase fünf bis sechs Nucleotide hinter der "Pribnow Box" (51). Die Purinbasen Guanin an Position -45 und Adenin an Position -34 können damit als erste Base des Transkripts fast ausgeschlossen werden, da die erstere sich gleich an die "TATA Box" anschließt, während die andere zu weit von dieser entfernt liegt.

Promotorregionen sind meist mit A-T reichen Abschnitten assoziiert (53). Aus der in Abbildung 7

gezeigten Verteilung des relativen Gehaltes an A-T und G-C Nucleotiden im pho A-Gen sowie seinen angrenzenden Abschnitten "downstream" und "upstream" geht deutlich hervor, daß der Anteil an A-T Paaren in der Promotorregion sehr hoch ist. Er beträgt 68% im Vergleich zu 47% im Strukturgenen der alkalischen Phosphatase. Diese Sequenzen spielen vermutlich sowohl bei der Erkennung durch die RNA-Polymerase als auch bei dem nachfolgenden lokalen Schmelzen der DNA eine entscheidende Rolle. Da die Bindungskräfte zwischen A-T Basenpaaren am niedrigsten sind, wird ihre Trennung energetisch begünstigt und demzufolge die mRNA-Kettenelongation beschleunigt. Wenn sich die Stärke eines Promotors daran messen läßt, so sollte die besonders auffallende Anhäufung solcher Paarungen zwischen der "Pribnow Box" und dem Initiationscodon des pho A-Gens als ein erster Hinweis auf das Vorliegen eines starken Promotors gewertet werden (77).

Die "Pribnow Box" ist Teil einer symmetrischen Sequenz, wie in Abbildung 8 gezeigt. Diese Sequenzen können durch Ausstülpung von Einzelstrangabschnitten aus der regulären DNA-Doppelhelix heraus Schleifen bilden und Signalwirkung haben (69). Auch die "-35 Region" ist Bestandteil eines Palindroms, wie Kikuchi et al. (78) durch Sequenzanalyse des Hind III-Bam I-Fragmentes nachweisen konnten. Diese zweidimensionale Struktur scheint jedoch nicht für die Erkennung des Promotors durch die RNA-Polymerase von Bedeutung zu sein, denn die Untersuchungen verschiedener Promotorregionen haben gezeigt, daß etwa 35 Basenpaare ausreichen, um die mRNA-Synthese spezifisch zu initiieren (79). Da die Expression des pho A-Gens komplexen Regulationsmechanismen unterliegt, ist es eher

wahrscheinlich, daß dieser Sequenzabschnitt "upstream" von der als Promotor angegebenen Nucleotidfolge als Bindungsstelle für Regulatorproteine dient. Offensichtlich sind die Produkte der Gene *pho B* und *pho R* für die Aktivierung der Transkription erforderlich (80). Es sind mehrere andere Gene beschrieben worden, die ebenfalls an der Steuerung der Genaktivität von *pho A* beteiligt sind (81, 82). Bisher ist es jedoch noch nicht gelungen, diese Aktivatoren und möglicherweise auch Repressoren näher zu charakterisieren, so daß die für die Wechselwirkung notwendigen Nucleotidsequenzen nicht festgelegt werden können. Mit Hilfe von *in vitro* Transkriptionsversuchen und durch Herstellung von Promotormutanten sollte es jedoch möglich sein, die gesamte für die Regulation der Genexpression notwendige DNA-Sequenz des Promotors zu ermitteln.

4.3 Terminator

Nicht das Stopcodon induziert den Abbruch der mRNA-Synthese, vielmehr enthält die DNA am Ende ihres codogenen Stranges spezifische Sequenzen, die die RNA-Polymerase als Terminationssignale erkennt. Dabei handelt es sich um einen G-C reichen und einen A-T reichen Abschnitt, die aufeinanderfolgen und teilweise oder ganz invertierte Wiederholungen vorhergehender DNA-Bereiche darstellen. Damit ist die Möglichkeit zur intrasträngigen Basenpaarung gegeben. Dies scheint für den Terminationsprozeß notwendig zu sein, denn bei Abwesenheit solcher Strukturen wurde die DNA *in vitro* fortlaufend transkribiert (53). Es wurde postuliert, daß die mRNA selbst ihre Elongation stoppt, indem sie ein "hairpin" ausbildet, das

entweder in eine Lücke des Polymerase-DNA-Komplexes eingefügt oder direkt an das Enzym gebunden wird, so daß sich die RNA-Polymerase von der DNA löst (83).

Symmetrische Sequenzen konnten auch am Ende des pho A-Gens nachgewiesen werden. Sie umfassen einen Teil der geforderten G-C und A-T reichen Region (1440 bis 1455). Inwiefern diese nicht vollständig invertiert wiederholten Sequenzen eine Regulatorfunktion besitzen, bleibt noch zu klären. Die Existenz der in allen bisher untersuchten Terminatoren gefundenen Strukturmerkmale legt jedoch den Verdacht nahe, daß dieser sich über 42 Basenpaare ausdehnende DNA-Abschnitt am Ende des pho A-Gens eine Terminatorsequenz darstellt.

4.4 Aminosäuresequenz

Die alkalische Phosphatase war das erste periplasmatische Enzym, für das aufgrund verschiedener in vitro-Versuche eine Signalsequenz gefordert wurde (84). So hatten Inouye und Beckwith (22) als Produkt des pho A-Gens eine Polypeptidkette gefunden, die größer war als die aus dem Periplasma isolierte monomere Untereinheit der alkalischen Phosphatase. Dieses Ergebnis konnte von Smith et al. (85) bestätigt werden. Sie bewiesen mit ihren Experimenten gleichzeitig, daß die alkalische Phosphatase wie alle bisher untersuchten sekretorischen Proteine in membrangebundenen Polysomen synthetisiert wird. Da dieses Precursormolekül im Gegensatz zur alkalischen Phosphatase mit großer Affinität an Decylagarose gebunden wurde, mußte ein hoher Anteil an hydrophoben Aminosäuren in dem zusätzlichen Abschnitt vorliegen (22). Seine Lage

aber - ob amino- oder carboxyterminal - wurde nicht festgelegt.

Bei der Bestimmung der Nucleotidsequenz des pho A-Gens und der daraus abgeleiteten Aminosäuresequenz ergeben sich 21 Aminosäuren am Aminoterminal des Translationsproduktes, die sich durch ihren hydrophoben Charakter als Signalpeptid zu erkennen geben. Auch von Inouye et al. (86) wurden diese Aminosäuren als Signalsequenz der alkalischen Phosphatase beschrieben. Sie weichen in ihrer Anzahl und Art jedoch erheblich ab von der von Lazdunski et al. (87) ermittelten Aminosäurezusammensetzung des Signalpeptids der alkalischen Phosphatase. Diese Autoren haben eine Methode zur in vitro-Isolierung solcher Vorläuferpeptide entwickelt. Sie basiert auf der Anreicherung der Präproteine an der Cytoplasmamembran durch Procain, das den Transport durch die Membran verhindert. Die Abtrennung der Signalsequenz gelingt durch Stimulation der Signalpeptidase mit Hilfe nicht-ionischer Detergenzien. Somit lassen sich diese unterschiedlichen Angaben vielleicht auf Verunreinigungen der Präparation durch andere Proteine zurückführen. Zum anderen besteht aber auch die Möglichkeit der Aktivierung einer unspezifischen Protease, so daß nicht das Signalpeptid, sondern ein anderes, kurzes Peptid abgespalten und isoliert wurde. Dafür spricht, daß nur acht von fünfundzwanzig Aminosäureresten in dem nach diesem Verfahren gewonnenen Signalpeptid als hydrophob eingestuft werden können.

Für das Strukturprotein wurden bereits sowohl an seinem Carboxyl- als auch an seinem Aminoende Teilsequenzen bestimmt (88, 89). Die Aufklärung der Gesamtsequenz des Isozyms III gelang 1981

Bradshaw et al. (43). Die durch Translation des Nucleotidcodes des pho A-Gens festgelegte Aminosäuresequenz stimmt mit dieser weitgehend überein mit Ausnahme der Aminosäuren Asparagin in Position 37 und 57. Das Triplett GAT codiert die Aminosäure Asparaginsäure, was durch die von Kikuchi et al. (78) durchgeführte Analyse der den aminoterminalen Abschnitt der alkalischen Phosphatase codierenden Sequenz bestätigt wird. Auch Bridgen und Secher (88) fanden bei der Untersuchung des Proteins an Position 37 Asparaginsäure, an Position 35 jedoch Glutaminsäure anstelle von Glutamin. In beiden Fällen wird deutlich, daß die Aminosäuren Glutamin und Asparagin bei der Proteinsequenzierung nur schwer von ihren entsprechenden Säuren zu unterscheiden sind.

Die von Bridgen und Secher (88) an erster Stelle genannte Aminosäure ist Arginin, die das Isozym I der alkalischen Phosphatase kennzeichnet und identisch ist mit der aus dem Nucleotidcode resultierenden Aminosäure an Position 22. Das Gen enthält also die Information zur Bildung aller drei beschriebenen Isozyme, die sich nur in ihrem Aminoterminus voneinander unterscheiden (61). Ihre Entstehung läßt sich daher nur auf eine posttranskriptionale Modifizierung zurückführen, deren Ursache bisher nicht aufgeklärt werden konnte. Es sind eine Vielzahl von Möglichkeiten diskutiert worden (89); unter anderem wurde die Bildung multipler Formen mit unterschiedlichen Wachstumsbedingungen erklärt (60). Es wäre aber auch denkbar, daß die Signalpeptidase selbst für die Bildung der Isozyme verantwortlich ist, indem sie die Signalsequenz mal vor und mal hinter dieser Aminosäure abtrennt. Da diese Enzyme sich jedoch durch eine hohe Substratspezifität auszeichnen,

kann eher vermutet werden, daß eine Protease diesen Argininrest abspaltet (90).

4.5 Folgerung

Die Expression der genetischen Information der in einen Plasmidvektor inserierten Fremd-DNA hängt davon ab, ob dieses Fragment eigene von dem Wirtsbakterium erkennbare Transkriptions- und Translationssignale trägt. Da die Regulationselemente der Säugergene von der Bakterienzelle jedoch nicht zur Proteinsynthese verwendet werden können, ist es notwendig, das eukaryontische Gen an einen Promotor des Vektors anzuschließen, so daß seine Transkription von dort aus reguliert wird (91). Bei Prokaryonten reicht das Startcodon alleine aber nicht aus, um die erfolgreiche Translation der zugehörigen mRNA zu initiieren (55). Vielmehr benötigen diese zusätzlich eine drei bis zwölf Basenpaare umfassende Sequenz, die sogenannte Shine Dalgarno Sequenz, die als Bindungsstelle für Ribosomen dient (56).

Mit Hilfe der Nucleotidsequenzanalyse wurde die Lage des pho A-Gens innerhalb des Fragmentes Hind III-Xho I des Plasmids pSB 53 festgelegt. Anhand charakteristischer Sequenzabschnitte konnte die für die Bindung der RNA-Polymerase notwendige Funktionseinheit des Promotors sowie eine zur 16S rRNA komplementäre Nucleotidfolge identifiziert werden. Vermutlich sind weitere für die Bindung von Regulatorproteinen erforderlichen Nucleotidsequenzen in dem Abschnitt Hind III-Alu I lokalisiert. Um diese bakteriellen Kontrollsequenzen zur Expression von Fremd-DNA auszunutzen, muß das

interessierende Gen in das Strukturgen der alkalischen Phosphatase eingefügt werden, so daß die mRNA-Synthese dann der Regulation des induzierbaren Transkriptionssystems des *pho A*-Gens untersteht. Für die Insertion des Säugergens bietet sich die *Hpa* II-Schnittstelle an, deren Erkennungssequenz 5' CCGG 3' an der Nahtstelle zwischen Signalsequenz und Strukturprotein codierendem Abschnitt des *pho A*-Gens liegt. Damit könnte unter Umständen sogar der Transport des synthetisierten Hybridproteins in den periplasmatischen Raum induziert werden. Da die Signalsequenz selbst während oder nach der Sekretion abgespalten wird, böte sich gleichzeitig die Möglichkeit, das dem Fremdgen entsprechende Produkt ohne bakterielle Proteinanteile aus dem Periplasma zu isolieren. Ein Beweis für den Erfolg dieses Verfahrens steht jedoch noch aus, obwohl dies für den Transport des Ratten-Proinsulins berichtet worden ist (92).

Andererseits könnte das aminoternale prokaryontische Teilstück des Fusionsproteins an einem Methioninrest mit Bromcyan abgespalten werden. Dies bedeutet jedoch, daß der Strukturbereich des Fremdgens mit einem zusätzlichen Methioninrest beginnen muß. Die selektive Hydrolyse der Peptidbindung gelingt aber nur, wenn das gewünschte Protein selbst kein Methionin enthält. Ist dies aber der Fall, so besteht die Möglichkeit, Schnittstellen zu konstruieren, die normalerweise in Proteinen nicht vorkommen. (93) Eine solche Schnittstelle ist zum Beispiel die Aminosäurefolge Pro-X-Gly-Pro. Sie wird von einer Collagenase, der Clostridio-Peptidase A, erkannt und zwischen X und Gly aufgespalten. Die am aminoterminalen Ende des eukaryontischen Proteins dann verbleibenden Amino-

säuren Gly und Pro können mit einer Exopeptidase (PPDA) abgespalten werden.

Beide Verfahren sind mit einem beträchtlichen experimentellen Aufwand verbunden. Eine wesentlich einfachere Methode besteht darin, die Fremd-DNA hinter einem mikrobiellen Promotor unmittelbar an das Initiationscodon anzuschließen. Das entstehende Protein besitzt dann zwar einen bakteriellen Methioninrest; dieser wird jedoch oft auf natürliche Weise von der Bakterienzelle abgetrennt. Um das Säugergen an das Initiationscodon GTG des pho A-Gens anzufügen, bedarf es einer geeigneten Schnittstelle. Eine solche läßt sich aus der Nucleotidsequenz des Gens nicht ableiten. Dieses Ziel könnte aber erreicht werden, wenn das Plasmid mit der Restriktionsendonuclease Xho I geöffnet und die Nucleotide sukzessiv mit dem Enzym Bal 31 bis zu der entsprechenden Stelle abgebaut würden.

5. Zusammenfassung

Das Plasmid pSB 53 enthält das die alkalische Phosphatase von Escherichia coli K 12 codierende pho A-Gen in dem Abschnitt Hind III-Xho I. Seine exakte Position innerhalb dieses Fragmentes sowie die für die Initiation und Termination der mRNA-Synthese erforderlichen Sequenzbereiche wurden durch Nucleinsäuresequenzanalyse festgelegt.

Das zur Sequenzbestimmung angewandte Verfahren nach Maxam und Gilbert fordert umfangreiche Kenntnisse zur Anordnung der Restriktionsschnittstellen, um das DNA-Molekül gezielt fragmentieren zu

können. Für die Abschnitte Hind III-Eco RI und Eco RI-Xho I wurden Restriktionskarten angefertigt, die aus einem Vergleich der nach vollständiger und partieller Spaltung erhaltenen Bandenmuster resultieren.

Das Strukturgen beginnt 280 Nucleotide hinter der Hind III-Schnittstelle mit dem außergewöhnlichen Triplet GTG und erstreckt sich über 1416 Basen bis in das Fragment Eco RI-Xho I hinein, wo es mit dem Stopcodon TAA endet. Es determiniert die Präphosphatase, ein Precursormolekül des Enzyms mit einem Molekulargewicht von 49404, das sich durch einen zusätzlichen aminoterminalen Abschnitt von dem aus dem Periplasma zu isolierenden Protein unterscheidet. Dieses sogenannte Signalpeptid besteht aus 21 Aminosäuren, das aufgrund seiner Lage und seines hydrophoben Charakters identifiziert wurde. Das Strukturprotein setzt sich aus 450 Aminosäuren zusammen. Mit Hilfe der Restriktionsendonuclease Hpa II läßt sich die Signalsequenz von der das Strukturprotein codierenden Sequenz abtrennen, so daß eukaryontische oder auch Gene anderer Herkunft an dieser Stelle angefügt und unter dem Einfluß des induzierbaren Transkriptionssystems des pho A-Gens auch exprimiert werden können.

Die Leserichtung des pho A-Gens entspricht nicht der in der Genkarte des E. coli Chromosoms angegebenen, denn dieses DNA-Molekül wird innerhalb des Plasmids im Uhrzeigersinn transkribiert. Die Transkription wird wahrscheinlich mit der Purinbase Guanin 39 Nucleotide vor dem Proteinsyntheseinitiationscodon eingeleitet. Die sich anschließende Sequenz trägt die Information für die nicht translatierte "leader Sequenz" der mRNA. Sie ist

gekennzeichnet durch die neun Nucleotide vom Startcodon entfernt liegende Shine Dalgarno Sequenz 5' GGAG 3', die mit dem 3'-Ende der 16S rRNA hybridisieren kann und somit die Bindungsstelle des Transkripts für Ribosomen darstellt.

Der 40 Basenpaare umfassende Abschnitt vor dem Initiationsspunkt der Transkription ist Teil des Promotors. Es wurden zwei Sequenzfolgen gefunden, die aufgrund ihres Abstandes zueinander und ihrer Ähnlichkeit mit der aus einem Vergleich verschiedener Promotoren abgeleiteten Modellsequenz als die "-35-Region" und "Pribnow Box" identifiziert werden konnten.

Am Ende des Gens wurde 23 Nucleotide hinter dem Stopcodon eine G-C-reiche Sequenz entdeckt, der sich eine Folge von fünf Thymidylresten anschließt. Demzufolge sollte dieser Bereich der Termination der mRNA-Synthese dienen. Ein obligatorisch Rho-Faktor abhängiger Terminator kann damit fast ausgeschlossen werden.

Bei der mit Hilfe eines Computerprogrammes durchgeführten Analyse der Nucleinsäuresequenz des pho A-Gens wurden alle Erkennungssequenzen für Restriktionsendonucleasen ermittelt, so daß ein Vergleich mit den experimentell bestimmten Daten möglich war. Zudem wurde die Sequenz systematisch nach der Häufigkeit des Auftretens von Codewörtern und invertiert repetitierten Sequenzen durchsucht. Dabei zeigten sich größere Symmetrieblocks nur in den für die Regulation der Transkription notwendigen Promotor- und Terminatorregionen. Aus der Verteilung der Basen über das gesamte Gen ergab sich ein erhöhter Anteil von G-C Paarungen, wäh-

rend in der 80 Basen langen Sequenz, die dem Gen vorgeschaltet ist, A-T Nucleotidpaare überwiegen, die einen starken Promotor kennzeichnen.

Literatur

1. Arber, W., *Ann. Rev. Microbiol.* 19, 365-378 (1965)
2. Smith, H.O., Wilcox, K.W., *J. Mol. Biol.* 51, 379 - 391 (1970)
3. Cozzarelli, N.R., Melechen, N.E., Jovin, T.M., Kornberg, A., *Biochem. Biophys. Res. Commun.* 28, 578-586 (1967)
4. Morrow, J.F., Cohen, S.N., Chang, A.C.Y., Boyer, H.W., Goodman, H.M., Helling, R.B., *Proc. Natl. Acad. Sci. USA* 71, 1743-1747 (1974)
5. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S.P., Chick, W.L., Gilbert, W., *Proc. Natl. Acad. Sci. USA* 75, 3727-3731 (1978)
6. Helinski, D.R., Clewell, D.B., *Ann. Rev. Biochem.* 40, 899-941 (1971)
7. Timmis, K.N., Cohen, S.N., Cabello, F.C., *Progr. Mol. Subcell. Biol.* 6, 1-58 (1978)
8. Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., Riggs, A.D., *Proc. Natl. Acad. Sci. USA* 65, 106-110 (1979)
9. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K., Weissmann, C., *Nature* 284, 316-320 (1980)
10. Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., Boyer, H.W., *Science* 198, 1056-1063 (1977)
11. Fraser, T.H., Bruce, B.J., *Proc. Natl. Acad. Sci. USA* 75, 5936-5940 (1978)
12. Martial, J.A., Hallewell, R.A., Baxter, J.D., Goodman, H.M., *Science* 205, 602-607 (1979)

13. Emtage, J.S., Tacon, W., Catlin, G.M., Jenkins, B., Porter, A.G., Carey, N.H., *Nature* 283, 171-174 (1980)
14. Ehrlich, S.D., *Proc. Natl. Acad. Sci. USA* 75, 1433-1436 (1978)
15. Kreft, J., Bernhard, K., Goebel, W., *Mol. Gen. Genet.* 162, 59-67 (1978)
16. Niaudet, B., Ehrlich, S.D., *Plasmid* 2, 48-58 (1979)
17. Hinnen, A., Hicks, J.B., Fink, G.R., *Proc. Natl. Acad. Sci. USA* 75, 1929-1933 (1978)
18. Siewert, G., Boidol, W., Daum, J., "Plasmidvektoren, Plasmide mit Genen für alkalische Phosphatasen und Bakterien, die letztere enthalten, ihre Herstellung und Verwendung". EP Nr. 0023882 (1981)
19. Inouye, H., Michaelis, S., Wright, A., Beckwith, J., *J. Bacteriol.* 146, 668-675 (1981)
20. Yoda, K., Kikuchi, Y., Yamasaki, M., Tamura, G., *Agric. Biol. Chem.* 44, 1213-1214 (1980)
21. Bachmann, B.J., Low, K.B., *Mikrobiol. Rev.* 44, 1-56 (1980)
22. Inouye, H., Beckwith, J., *Proc. Natl. Acad. Sci. USA* 74, 1440-1444 (1977)
23. Blobel, G., Dobberstein, B., *J. Cell. Biol.* 67, 835-851 (1975)
24. Clewell, D.B., Helinski, D.R., *Proc. Natl. Acad. Sci. USA* 62, 1159-1166 (1969)
25. Humphreys, G.O., Willshaw, G.A., Anderson, E.S., *Biochem. Biophys. Acta.* 383, 457-463 (1975)
26. Southern, E., in "Methods in Enzymology" (Wu, R., Ed.) Academic Press, New York, Vol. 68, 152-176 (1979)
27. Hayword, G.S., *Proc. Natl. Acad. Sci. USA* 71, 2108-2112 (1974)

28. Blin, N., Gabain, A.V., Bujard, H., FEBS Lett. 53, 84-86 (1975)
29. Maxam, A.M., Gilbert, W., in "Methods in Enzymology" (Grossmann, L., Moldave, K., Eds.) Academic Press, New York, Vol. 65, 499-560 (1980)
30. Smith, D.R., Calvo, J.M., Nucl. Acids Res. 8, 2255-2274 (1980)
31. Lillehaug, J.R., Kleppe, R.K., Kleppe, K., Biochemistry 15, 1858-1865 (1976)
32. Tu, C-P.D., Cohen, S.N., Gene 10, 177-183 (1980)
33. Instruction manual for labelling the 3'end of nucleic acid der Firma NEN
34. Smith, H.O., Birnstiel, M.L., Nucl. Acids Res. 3, 2387-2398 (1976)
35. Burton, K., in "Methods in Enzymology" (Grossmann, L., Moldave, K., Eds.) Academic Press, New York, Vol. 12 A, 222-225 (1967)
36. Peacock, A.C., Dingman, C.W., Biochemistry 6, 1818-1827 (1967)
37. Isono, K., Nucl. Acids Res. 10, 85-89 (1982)
38. Szalay, A.A., Grohmann, K., Sinsheimer, R.L., Nucl. Acids Res. 5, 1569-1578 (1977)
39. Ohmori, H., Tomizawa, J., Maxam, A.M., Nucl. Acids Res. 5, 1479-1485 (1978)
40. Boidol, W., Simonis, M., Töpert, M., Siewert, G., Mol. Gen. Genet. 185, 510-512 (1982)
41. Sarthy, A., Fowler, A., Zabin, I., Beckwith, J., J. Bacteriol. 139, 932-939 (1979)
42. Bradshaw, R.A., Cancedda, F., Ericsson, L.H., Newmann P.A., Piccoli, S.P., Schlesinger, M.J., Shrieffer, K., Walsh, K.A., Proc. Natl. Acad. Sci. USA 78, 3478-3477 (1981)
43. Miller, J.H., Cell 1, 73-76 (1974)
44. Sarthy, A., Michaelis, S., Beckwith, J., J. Bacteriol. 145, 293-298 (1981)

45. Wanner, B.L., Wieder, S., Mc Sharry, R., J. Bacteriol. 146, 93-101 (1981)
46. Brown, N.L., Smith, M., J. Mol. Biol. 116, 1-30 (1977)
47. Epstein, W., Beckwith, J.R., Ann. Rev. Biochem. 37, 411-436 (1968)
48. Rosenberg, M., Court, D., Ann. Rev. Genet. 13, 319-353 (1979)
49. Pribnow, D., Proc. Natl. Acad. Sci. USA 72, 784-788 (1975)
50. Schaller, H., Gray, C., Herrmann, K., Proc. Natl. Acad. Sci. USA 72, 737-741 (1975)
51. Scherer, G.E.F., Walkinshaw, M.D., Arnott, S., Nucl. Acids Res. 5, 3759-3773 (1978)
52. Kleid, D., Humayan, Z., Jeffrey, A., Ptashne, M., Proc. Natl. Acad. Sci. USA 73, 293-297 (1976)
53. Pribnow, D., in "Biological Regulation and Development" (Goldberger, R.F., Ed.) Vol. I, 219-277 (1978)
54. Walz, A., Pirrotta, V., Ineichen, K., Nature 262, 665-669 (1976)
55. Steitz, J.A., Jakes, K., Proc. Natl. Acad. Sci. USA 72, 4734-4738 (1975)
56. Shine, J., Dalgarno, L., Proc. Natl. Acad. Sci. USA 71, 1342-1346 (1974)
57. Ito, K., Bassford, P.J., Beckwith, J., Cell 25, 143-150 (1981)
58. Devillers-Thiery, A., Kindt, T., Scheel, G., Blobel, G., Proc. Natl. Acad. Sci. USA 72, 5016-5020 (1975)
59. Randall, L.L., Hardy, J.S., Josefsson L.-G., Proc. Natl. Acad. Sci. USA 75, 1209-1212 (1978)
60. Natori, S., Garen, A., J. Mol. Biol. 49, 577-588 (1970)

61. Schlesinger, M.J., Block, W., Kelly, P.M., in "Isozymes I: Molecular Structure" (Markert, C., Ed.) Academic Press, New York, 333-342 (1975)
62. Nomura, M., Post, L.E., Jinks, S., in "RNA Polymerase, tRNA and Ribosomes" (Osawa, S., Ed.) University of Tokyo Press, Japan, 315-328 (1980)
63. Air, G.M., Blackburn, E.H., Sanger, F., Coulson, A.R., J. Mol. Biol. 96, 703-719 (1975)
64. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G., Ysebaert, M., Nature 260, 273-278 (1976)
65. Sanger, F., Air, G.M., Barell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C. Hutchison III, C.A., Slocombe, P.M., Smith, M., Nature 265, 687-695 (1977)
66. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M., Mercier, R., Nucl. Acids Res. 9, r43-r74 (1981)
67. Ikemura, T., J. Mol. Biol. 146, 1-21 (1981)
68. Dykes, G., Bambara, R., Marians, K., Wu, R., Nucl. Acids Res. 2, 327-345 (1975)
69. Gierer, A., Nature 212, 1480-1481 (1966)
70. Smith, H.O., Science 205, 455-462 (1979)
71. Yuan, R., Ann. Rev. Biochem. 50, 285-315 (1981)
72. Sato, S., Nakazawa, K., Shinomiya, T., J. Biochem. 88, 737-747 (1980)
73. Roberts, R.J., Nucl. Acids Res. 9, r75-r96 (1981)
74. Mann, R.A., Smith, H.O., Nucl. Acids Res. 4, 4211-4221 (1977)
75. Rubin, R.A., Modrich, P.J., J. Biol. Chem. 252, 7265-7272 (1977)

76. Dickson, R.C., Abelson, J., Barnes, W.M., Reznikoff, W.S., *Science* 187, 27-35 (1975)
77. Chamberlin, M.J., in "RNA Polymerase" (Losick, R., Chamberlin, M., Eds.) Cold Spring Harbour Laboratory, New York, 159-191 (1976)
78. Kikuchi, Y., Yoda, K., Yamasaki, M., Tamura, G., *Nucl. Acids Res.* 9, 5671-5678 (1981)
79. Khorana, H.G., Agarwal, K.L., Besmer, P., Büchi, H., Caruthers, M.H., Cashion, P.J., Fridkin, M., Jay, E., Kleppe, K., Kleppe, R., Kumar, A., Loewen, P.C., Miller, R.C., Minamoto, K., Panet, A., Rajbhandary, U.L., Ramamoorthy, B., Sekiya, T., Takeya, T., van de Sande, J.H., *J. Biol. Chem.* 251, 565-570 (1976)
80. Echols, H., Garen, A., Garen, S., Torriani, A., *J. Mol. Biol.* 3, 425-438 (1961)
81. Willsky, G.R., Bennett, R.L., Malamy, H., *J. Bacteriol.* 113, 529-539 (1979)
82. Wanner, B.L., Sarthy, A., Beckwith, J., *J. Bacteriol.* 140, 229-239 (1979)
83. Adhaya, S., Gottesman, M., *Ann. Rev. Biochem.* 47, 967-996 (1978)
84. Halegoua, S., Inouye, M., in "Bacterial Outer Membranes, Biogenesis and Functions" (Inouye, M., Ed.) John Wiley and Sons, New York, 67-113 (1979)
85. Smith, W.P., Tai, P.C., Thomson, R.C., Davis, B.D., *Proc. Natl. Acad. Sci. USA* 74, 2830-2834 (1977)
86. Inouye, H., Barnes, W., Beckwith, J., *J. Bacteriol.* 149, 434-439 (1982)
87. Lazdunski, C., Baty, D., Pagès, J.-M., *Eur. J. Biochem.* 96, 49-57 (1979)
88. Bridgen, J., Secher, D.S., *FEBS Lett.* 29, 55-57 (1973)

89. Kelley, P.M., Neumann, P.A., Shriefer, K., Cancedda, F., Schlesinger, M.J., Bradshaw, R.A., *Biochemistry* 12, 3499-3503 (1973)
90. Piggot, P.J., Sklar, M.D., Gorini, L., J. *Bacteriol.* 110, 291-299 (1972)
91. Guarente, L., Roberts, T.M., Ptashne, M., *Science* 209, 1428-1430 (1980)
92. Talmadge, K., Kaufman, J., Gilbert, W., *Proc. Natl. Acad. Sci. USA* 77, 3988-3992 (1980)
93. Töpert, M., Dumann, L., Matthes, S., Lübke, K., Büllersbach, E., Abstract of Communications, 14th FEBS Meeting, Edinburgh 1981, *Biochem. Soc. Trans.* 9, 272 (1981)

Lebenslauf

Zur Person: Marianne Simonis

geb. am 14.6.1955 in Karl/Eifel

Eltern: Karl Simonis

Angelika Simonis geb.

Barzen

Pflege-

eltern:

Prof. Dr. Wolfgang Simonis

Irmgard Simonis, geb.

Lainck-Vissing

Schulzeit: 1962-1966

Grundschule in Karl

1966-1974

neusprachliches

Mädchengymnasium in

Gladbeck

Juni 1974

Reifeprüfung

Studienzeit: WS 1974/75

Aufnahme des Studiums der

Biologie an der FU Berlin

Januar 1977

Diplom-Vorprüfung

WS 1977 -

Tätigkeit als studentischer

SS 1980

Tutor am Institut für

Biophysik

Dezember

1978

Diplomprüfung

SS 1979 -

Anfertigung der Diplom-

SS 1980

arbeit "Kinetische Charak-

terisierung der membran-

gebundenen Nucleotidase

von Meeresbakterien (MB 22)

am Institut für Biochemie

und Molekularbiologie

WS 1980/81

Beginn der Dissertation in

den Forschungslaboratorien

der Schering AG, Berlin

